

FIGURE 1

GGGGCTCGGCCAGCGCAGCGTAGTCGGTCTGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGTAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCCCTCAGCCCTGTAATTGGACATCTGCTGCTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATAAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGCAGTTTATGCATTG
CTACCATTTATGTCGTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGCCCTGTACTTGGAAACTGAGTTGTTAGGACTTCTATTGTGGCAAACCTT
CCAGAAAAACAACCCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCTCAT
TATATATGTTGTTCAAGACATCCTTACCAAAATGCAGCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGGTTATCTGGTGTGGAGTAAGTCACCTAGCATGCTGACTTGCTC
ATCAGTTTGACAGTGGCAATTGGGACTGATTAGAACAGAAACTCCATTGGAACCCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTCTTCTT
GGTTTTCTGACTTACATTGCTGATTTCAGAAAATTCTTACGGGTGGAAGCCAATTACA
TGGATTAACCCCTATGACACTGCACCTGCCCTATTAACAATGAACGAACACGGCTACTTCCA
GAGATATTTGATGAAAGGATAAAATTTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGAAATTTCACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAACGCCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAACACCTATGCCTATACCTTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYTAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLHQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACCGTGGCGGACCGTGGGGAGAGCCGCAGTCGGCTGCAGCACCTGGAGAAGGCAGACC
GTGTAGGGGGCCTGCCCCAGCGTGTGGCTCGGGAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCTGAGTTTCCCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTG
GATTGGGTGGCTTTCTTCATCGCCAATTGTTAAAGACTATGAGATACTGAGTATGTTGACAG
GTGATCTCTCGTGTGACGTTGCATTTCTGCACCATGTTGAGCTCATCATCTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCGTTATTCACTGGAAAATGAACCTGTGTGAAATTCTGCTGATCTGG
TTTCATGGTGCCTTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCATAAACACGA
CTGCTTTTCTGCTCTTATGGTGACCTTATGTATTCTCTGAAACTAGGAGATCCCTTCC
CATTCTCAGCCAAAACATGGATCTTATCCATAGAACAGCTCATCAGCCGGTTGGTGTGATTGGAG
TGAECTCTCATGGCTCTTCTGGATTGGTGTCACTGCCATACACTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATTCTAGCCCTGGAACGGCAGTGTCAAACCATGGATATGAT
CATAAAGAAAAAGAAAAGGATGGCAATGGCACGGAGAACAAATGTTCCAGAAGGGGAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAAGTGTACCACTTCAGCATCAGGAAGTGAATCTACT
CTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAAGCAGGCAGCTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAACCTCAAGGGAAATATTTAATTTCTTGGTT
ACTTTTCTCTATTACTGTGTTGGAAAATTTCATGGCTACCACATCAATATTGTTTGATCGAGTT
GGGAAAACGGATCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGAAATCCAATTGATGT
GAAGTTTGGTCCAACACATTCTTCTCATCTTGTGGAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTATGCCATCTCTAGCAGTAAGTCCTCAATGTCATTGCTCG
CTATTAGCACAGATAATGGCATGTACTTGTCTCTGTGCTGATCCGAATGAGTATGCCTT
AGAATACCGCACCATAAATCACTGAAGTCCTGGAGAACACTGCAGTTCAACTTCTATCACCGTGGTTG
ATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCTCTATGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAAATTAA
GATATAAGAGGGGGAAAATGGAACCAGGGCTGACATTATAAACAAACAAAATGCTATGGTAGC
ATTTTCACCTTCATAGCATACTCCTCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGAGAGAACACTAAGACAATACTCAGCAGAGAGCATCCGTGTGGATATGAGGCTGG
TGTAGAGGCCAGAGGAGCAAGAAAATGAGTAAAGGTTACATGGAACTCTGGCAAGACATGT
CTATGGTAGCTGAGCCAACACGCTAGGATTCCGTTAAGGTTACATGGAAAGGTTAGCTTTG
CCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAAAAAAAAGGGCGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTGGCCCATGGCCAACCTGTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFFWKLGDP
FPILSPKKGILSIEQLISRGGVGIVGVTIMALLSGFGAVNCPTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALEELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVKGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGMY
FVSSVLLIRMSMPLYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCCATCCGACTAGTTATTGAGCATCTGCCTCTATACACCAGTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGATGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
TCCATCTGGACCACGAGGCTCTGGTCCAAGGCTCTTGCCTGAGAAGAGCTTCCATCCAGGT
GTCATGCAGAATTATGGGATCACCCCTGTGAGAAAAAGGGGAACCAGCAGCTGAATTACAG
AAGCTAAGGAGGCCGTAGGCTGGGACTAAGTTGGCCGGCAAGGACCAAGTGAAACAGCC
TTGAAAGCTAGCTTGAACACTGCAGCTATGGCTGGGGATGGATTGCTGGTCACTCTAG
GATTAGCCCCAACCCCAAGTGTGGGGAAATGGGGTGGGTGCTGTGTTGAAAGGTTCCAGTGA
GCCGACAGTTGCAGCTATTGTTACAACACTCATCTGATACTTGGACTAACTGTGCATTCCAGAA
ATTATCACCAACAAAGATCCCATATTCAACACTCAAACGTCAACACAAACAGAATTATTGT
CGTGCAGTACCTACTCGGTTGCATCCCCCTACTCTACAATACCTGCCCTACTACTCCTC
CTGCTCCAGCTTCACTTCTATTCCACGGAGAAAAAAATTGATTGTGTCACAGAAGTTTATG
GAAACTAGCACCATGTCTACAGAAACTGAACCATTTGGTAAAGAAAGCAGCATTCAAGAATGA
AGCTGCTGGGTTGGAGGTGCCCCACGGCTGCTAGTGTGCTCTCCCTCTTGGTGTG
CAGCTGGTCTGGATTGCTATGTCAGGATTTGAGGTTACAAACAGAAT
CAGCAGAAGGAAATGATCGAACACAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTAA
TGAGGAATCAAAGAAAATGATAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACCTACCGTGC
GATGCTTGAAGCTGAAGTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTT
CATGCTCTTACCTGCCCCAGCTGGGAAATCAAAGGGCCAAAGAACCAAAGAACAGAAATCCA
CCCTGGTTCTTAACGGAAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAT
GCCCTCTCTTATTGTAACCTGTCTGGATCTTACCTCTACCTCCAAAGCTCCCACGGCT
TTCTAGCCTGGCTATGCTCTAATAATATCCCACGGAGAAAGGAGTTTGCAAGTGCAGGAC
CTAAACATCTCATCAGTATCCAGTGGAAAAAGGCCCTCTGGTGTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTACTGATTCCGCAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTCTGAGGCCGGTAAGAGCAAAGAAT
GGCAGAAAAGTTAGGCTGAGGACATGGAGATTCTCATACTGAGACCTAATCTCTGTAAA
GCTAAAATAAAGAAATAGAACAGGCTGAGGACATGAGCAGTACACTGTCAGCAGGGACTGTAAAC
ACAGACAGGGTCAAAGTGTCTCTGAAACACATTGAGTTGAACACTGTTAGAACACACAC
CTTACTTTCTGGTCTCTACCAACTGCTGATATTCTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAAATCTTATAAATTCTATTCTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCTCTAAAAAATTGCACATAGTAG
AACGCTATCTGGGAAGCTATTCTCTGAGTTGATATTCTAGCTTATCTACTTCCAAACATAAT
TTTATTTCTGCTGAGACTAATCTTATTCTCATATGGCAACCATTATAACCTTAATT
TATTATTAACACCTAACAGTACATTGTTACCTATATACCAAGCACATTAAAAGTGC
ATTAACAAATGTATCACTAGCCCTCTTTCAACAGAACAGGACTGAGAGATGCAGAAATT
TGTGACAAAAAATTAAAGCATTAGAAAACCTT

FIGURE 6

MARCFSLVLLLTSIWTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNNPKCGKNGVGVLIWKPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAASKNEAAGFGGVPALLVLALLFFGAAAGLGFCYVK
RYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGACCCGGGCCCGCCACCAGCGCCGCTCCCGCATCTGCACCCGAGCCCGC
GGCCTCCGGCGGGAGCGAGCAGATCCACTCCGCCCCGAGCGAACCTCGGTCCAGTCGGCG
CGGCTGCGGGCGCAGAGCGGAGATGCCAGCGGCTTGGGCCACCCCTGCTGTGCCCTGCTGGCG
CGGCGGTCCCCACGGCCCCCGCGCCGCTCGACGGCACCTCGGTCCAGTCAGCAAGCCGGCCG
GCTCTCAGCTACCCGAGGAGGAGGCCACCTCAATGAGATGTTCCCGAGGTTGAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGTGGAAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGAAATAATACCATCCATGTGCACCGAGAAATTACAAGATAACCAACAACCAGAC
TGGACAAATGGTCTTTCAGAGACAGTTATCACATCTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTAUTGCCAGTTGCCAGCTCCAGTAC
ACCTGCCAGCCATGCCGGGGCAGAGGATGCTCTGCACCCGGGACAGTGAGTGTGTGGAGACCA
GCTGTGTCTGGGTCAGTGCACCAAAATGGCCACCAGGGCAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCGGGCTGTGTGCTTCCAGAGAGGCCTGCTGTTCCGTG
ACACCCCTGCCGTGGAGGGCGAGCTTGCATGACCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGGCTGATGGAGCCTGGACCGATGCCCTGTGCAAGGCTGCTGTTCCGTG
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTCGTGGGAGCCGTGACCAAGATGG
GAGATCTGCTGCCAGAGAGGGTCCCGATGAGTATGAAGTTGGCAGCTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCCTGGGGAGCCTGCGCTG
CCGCCGCTGACTGCTGGGAGGGAGAGATTAGATCTGGACCAAGGCTGTGGTAGATGTGCAA
TAGAAATAGCTATTATTCAGGTGTGCTTTAGGCCTGGCTGACCAGGTTCTCCTA
CATCTTCTCCCAGTAAGTTCCCTCTGGCTTGACAGCATGAGGTGTTGCATTGTTAGCT
CCCCCAGGCTGTTCTCAGGCTTACAGTCTGGCTGGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTGCCACCCCTGTCAGATTATTGGCTGCTTGCCTCTACCAAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTGATAATTGTTGAGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC
TGATTGGTTGGGAAATGTGGAGAAGAGTGCCTGTTGCAAACATCAACCTGGAAAATG
CAACAAATGAATTTCACGCAAGTTCTTCATGGCATAGGTAAGCTGTGCCTCAGCTGTTGC
AGATGAATGTTCTGTTACCCCTGCATTACATGTGTTATTCATCCAGCAGTGTGCTCAGCTCC
TACCTCTGCCAGGGCAGCATTTCATATCCAAGATACTCCCTCTCAGCACAGCCTGGGG
AGGGGGCATTGTTCTCTCGCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTGTGACTCTAAAGCTCAGTGTCT
CTCCACTACCCACACCAGCCTGGTGCACCAAAAGTGCTCCAAAAGGAAGGAGAATGGGAT
TTTCTGAGGATGACATCTGGAAATTAGGTCAAACATAATTCTCACATCCCTCTAAAGTAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGAGCCGTCTTAATGAAGACAATGAT
ATTGACACTGTCCTTGGCAGTTGCATTAGTAACCTGAAAGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAAGAACAGTACTTAGGTAATTGTTAGGGCAGGATTATAATGAAATTG
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGAATATGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGATG
TTTCAGGTGTCTGGACTGTTGCCACCATGTATTCTCAGAGTTCTTAAAGTTAAAGTTGCA
CATGATTGTATAAGCATGCTTCTTTGAGTTAAATTATGTATAAACATAAGTTGCATTAGAA
ATCAAGCATAAAATCACTCAACTGCAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNNTIHVHREIHKITNNQTGQMVFSE
TVITSGVDEEGRRSHECIIDECDGPSMYCQFASFQYTCQPCRGRMLCTRSECCGDQLCVWGHC
TKMATRGSGNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEAAAAALLGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACCGTGGCGGACCGTGGGGCTGTGAGAAAGTCCAATAAACATCATGCAACCCAC
GGCCCACCTTGTGAACCTCTCGGCCAGGGCTGATGTGCGTCTTCAGGGCTACTCATCAAAG
GCCTAATCCAACGTTCTGTCTCAATCTGAAATCTATGGGTCTGGGCTCTCTGGACCTT
AACTGGGTACTGGCCCTGGCCAATGCGTCTCGCTGGAGCCTTGCCCTCTACTGGCCTT
CCACAAGCCCCAGGACATCCCTACCTCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTATTGGCATTGGAGGCCCTACCTGACCCCTGTGCAGATAGCCGGGTATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCTGTAGCCGCTGCATCATGTGCTGTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTATCAAGTTCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTCTGTCTCAGCCAAAATGCGTTCATGCTACTCATGCGAAC
ATTGTCAGGGTGGTCGTCTGGACAAAGTCACAGACCTGCTGCTGTTCTTGGGAAGCTGCTGGT
GGTCGGAGGCCTGGGGGCTCTGCTCTCTCTCCGGTCGATCCGGGCTGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCCATGACCTCCATCCTGGGGCTAT
GTCATCGCCAGCGGCTCTTCAGCGTTTCGGCATGTGTGGACAGCCTCTCCCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCAGGCTACTACATGTCCAAGAGCCTCTAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCTTACAGGT
CTCCATTGTTGGTAAAAAAAGGTTTAGGCCAGGCGCGTGGCTACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCAGGCGGATCACCTGAGTCAGGAGTTGAGGACCCAGCCTGCCAACATGGTG
AAACCTCCGTCTCTATTAAAAACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCACTCCA
GCTACTCGGGAGGTGAGGCAGGAGAACGCTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAA
AAGATTTATTAAAGATTTTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCLWCLEKFIKFLRNAYIMIAIYGKNFCVSAKNAFMILLMRN
IVRVVVLDKVTDLLLFFGKLLVVGGVGVLSSFFFFSGRIPGLGKDFKSPHLNYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNNGSLDRPYYMSKSLLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

CCCCCGCGCCCGGCCGGCGCCGGCGCCGAAGCCGGGAGCCACCGCCATGGGGGCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCTGCCCTCTGCCGTCTGCCCTGCATCCGTGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCTCATCTTCACGTTCTCCCTCCCTGGGTGCTGGTGTCCA
TCATTATGCTGAGCCC GGCGTGGAGAGTCAGCTACAAGCTGCCCTGGGTGTTGAGGAGGGGCC
GGGATCCCCACCGTCTGCAGGCCACATCGACTGTGGCTCCCTGCTGGCTACCGCGTGTCTACCG
CATGTGCTCGCCACGGCGCCTCTCTCTTTCACCTGCTCATGCTTGCGTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTTGGTCTTTAAGTCTGATCCTGGTGGCCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCCTCACCAACATCTGGTTCTACTTCGGCGTCGTGGG
CTCCTCCCTCTTCATCCTCATCCAGCTGGTGCTGCTCATCGACTTGCACACTCTGGAACCGCGGT
GGCTGGGCAAGGCCAGGGAGTGCAGTCCCGTGCTGGTACGCAGGCCCTCTCTTCACTCCCTC
TTCTACTTGCTGTCGATCGCGCCGTGGCGTGATGTTCATGTA CACTGAGCCCAGGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCTCACCTCACCTCTGTGTCTGCGTGTCCATGCTGCTGTCTGC
CCAAGGTCCAGGACGCCAGGCCAACCTCGGGTCTGCTGCAGGCCCTCGGTATCACCTCTACACCAGT
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCAACCCAGCT
GGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGGATGCCCGAGCATTG
TGGGCCCTCATCATCTCCCTCTGTGCA CCTCTTCATCAGTCTGCGCTCCTCAGACCACGGCAGGTG
AACAGCCTGATGCA GACCGAGGAGTGCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGT
GGCAGCCTGTGAGGCCGGGCTTGAACAGCAGCAGGCCGTACCTACAGCTACTCCCTTCC
ACTTCTGCCCTGGTGTGGCCTACTGCACGT CATGATGACGCTCACCAACTGGTACAAGCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGTGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCTGCCAACCGCGACTTCAGCTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCTGCCACCTGGTGCCCTCTGGCTCGGTGACAGCCAACCTGCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGGC
CTTCTAGCGTAGTGCCTCAGGTCCGAGGAGCATCAGGCTCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCCTTCCCTCCCTCCCTGTTGCCCATACTCAGCATCTGGATGAA
AGGGCTCCCTGTCCCTCAGGCTCCACGGGAGCAGGGCTGCTGGAGAGAGCAGGGGAACTCCCACCAAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTGGTCAGTCCCCAGGGACCCCTGCCCTTCC
GACTTCGTGCCTTACTGAGTCTAAGACTTTCTAATAAACAAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHAWNQRWLKAE
ECDSRAWYAGLFFFLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIIFLLCTLFISLRSSDHRQVNLSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWVKICASWAGLLLWTLVAPLLLNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCAGGAACCACCGTTAAGGTGTCCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCATGCTCTGC~~CA~~ATATCCATTCCATCAACCCCACACAA
CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTGTTGTCACCTTGACCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTCTCTGGCAGTTTCAAGTGTAAACTTGCAATGCTGTGAG
ACTGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAG
TGATCCTTCGAAGCTTTCTCAAGGGCTTTGGCTATGTGCTGCCATCATTCAATT
CTTGCCTGGATTGAGACGTGGTCTGGATTCAAAGTGTACCTCAAGAACAGAAGAAGAAAA
CAGACTCCTGATAAGTCAGGATGCTCAGAGAGGGCAGCACTTACCTGGTGGCTTCTGATG
GTCAGTTTATTCCCCCTCTGAATCCGAAGCAGGATCTGAAGAACGTAAGAAAAACAGGACAGT
GAGAAACCACTTTAGAACTATGAGTACTACTTTGTTAAATGTGAAAACCC~~T~~CACAGAAAGTC
ATCGAGGCAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTACGTC
CACTGCTGGTTATTGAACAGCTAATAAGATTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTGCCTGCCTGGCTGGTAAGGTAATGTCATGATT~~C~~ATCCTCT
TCAGTGAGACTGAGCCTGATGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTAGTAAGCAAGATAACCTTTATTCAATT
AGAATGGAATTTTTGTTCATGTCAGATTATTTGTTATTCTTTAACACTCTACATT
TCCCTGTTTTAACTCATGCACATGTGCTTTGTCAGTTAAAGTGTAAATAAAATCTG
ACATGTCAATGTGGCTAGTTTATTCTGTTGCATTATGTTGATGGCCTGAAGTGTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAGTCAACAGACATTGTATT
TTATCATGAAATCATGTTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTGAATGC
ACAAAATGACTTAAACCATT~~C~~ATATCATGTTCCCTTGC~~G~~TTCAAGCCAATTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWEITWFLDFKVLQPQEAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCCTGGGGCCGCCAGGAAAGACTGAGG
CCGCGGCCTGCCCCCGCCGCTCCCTGCGCCGCCCTCCCGGGACAGAAGATGTGCTCCAG
GGTCCCCTGCTGCTGCTGCTCCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGCTGCCCAT
CCGGCTGCCAGTGAGCCAGCACAGACAGTCTCTGCACTGCCGCCAGGGACCACGGTCCCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTGAAGAACGGCATCACCATGCTGACGC
AGGCAGCTTGGGGCTGCCCCGCTGAGCTCTGACCTGTACAGAACAGATGCCAGCC
TGCCCAAGGGTCTTCCAGCCACTGCCAACCTCAGCAACCTGGACCTGACGCCAACAGGCTG
CATGAAATACCAAATGAGACCTTCCGTGGCTGCGGCCCTGAGCGCCTCACCTGGCAAGAA
CCGCATCCGCCACATCCAGCTGGTGCCTCGACACGCTGACCGCCTCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCCGGCACTGCCCGCTGCGCTGCCGCCCTGCTGCTGGACCTCAGC
CACAAACAGCCTCTGGCCCTGGAGGCCGATCTGGACACTGCCAACAGTGGAGGGCGCTGGGCT
GGCTGGTCTGGGGCTGCGACAGCTGGAGCGAGTGCACCTGTGATCCGAGGGCTCCGGGCTACG
TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCACCTGTGATCCGAGGGCTCCGGGCTGACG
CGCCTGCCGCTGGCCGAACACCCGATTGCCAGCTGCCGCCCCAGGACCTGGCGGCCCTGGC
TGCCCTGCCAGGAGCTGATGAGCAACCTAACGCTGCCAGGGCTGCCGACCTCTGGG
TCTTCCCCCGCTGCGCTGCTGGAGCTGCCAACCCCTTAACCTGCGTGTGCCCCCTGAGC
TGGTTGGCCCTGGTGCAGAGAGCCACGTACACTGCCAGCCCTGAGGAGACGCCGCTGCCA
CTTCCGCCAAGAACGCTGCCGGCTGCTCTGGAGCTGACTACGCCGACTTGGCTGCCAG
CCACCACCAACAGCACAGTGCCAACACAGGCCCCGGTGGTGGGGAGGCCACAGCCTTGCT
TCTAGCTGGCTCTACCTGGTAGCCCCACAGGCCGCAACTGAGGCCCCAGGCCGCCCTC
CACTGCCCAACCGACTGTAGGGCTGTCCCCAGGCCAGGACTGCCACCGTCCACCTGCC
ATGGGGCACATGCCACCTGGGACACGCCAACCTGGCGTGTGCTGTGCCAGGCTTCACG
GGCCTGTACTGTGAGAGGCCAGATGGGGCAGGGGACACGGCCAGGCCAACCCAGCAG
GCCACACGGCTCCGACCCCTGGGATCGAGCCGGTGAAGGCCACCTCCCTGCCGTGGGCTGC
AGCGCTACCTCCAGGGGAGCTCGTGCAGCTCAGGAGCCTCCGTCTCACCTATGCCAACCTATCG
GGCCCTGATAAGGGCTGGTGCAGCTGCCACTGCCCTGCCGTCTGAGTACACGGTCAACCC
GCTGCCGCCAACCCACTTACCGTCTGTGATGCCCTGCGGGGGGGGGGGGGGGGGGG
GCGAGGAGGCCGAGGGCAACCTGCCCTCATGGCCGCCCCCTGCGGGGGGGGGGGGGGG
CAGGCCGCGAGGGCAACCTGCCCTCATGGCCGCCCCCTGCGGGGGGGGGGGGGGGGG
GCTGGCTGCCGTGGGGGAGGCCACTGTGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
ACAAAGGGCAGGTGGGGCAGGGCTGGGGCCACTGGGAGGGGGGGGGGGGGGGGGGGGG
CCAGGGCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGGGGTCTGAGTGTGAGGTGCC
ACTCATGGCTTCCCAGGGCTGGCTCAGTCACCCCTCCACGCCAAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGCCGGCTCTCAGCCAGTGAAGATGCCAGCCCCCTCCTGCTGCC
ACACCACTGAAGTCTCAGTCCAAACCTGGGATGTGTGAGACAGGGCTGTGACCAAGCT
GGGCTGTCTGCCCTGGAACCTCGGTCTCTCATCTGTGAGATGCTGTGGGCCAGCTGACGCC
CTAACCGTCCCCAGAACCGAGTGCTATGAGGACAGTGTCCGCCCTGCCAACGTGCAAGTC
CCTGGGCACGGCGGGCCCTGCCATGTGCTGTAACGCATGCCCTGGGTCTGCTGGGCTCTCCAC
TCCAGGCCGACCCCTGGGGGCCAGTGAAGGAAGCTCCCCGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAAACTGAAAGGAAGATGC
TTTAGGAACATGTTTGTCTTTTAAATATATATATTATAAGAGATCCTTCCCATTATTCT
GGGAAGATGTTTCAAACCTCAGAGACAAGGACTTTGGTTGTAAAGACAAACGATGATGAA
GGCCTTGTAAAGAAAAATAAAAGATGAAGTGTGAA

FIGURE 16

MCSRVPLLLPLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDSLQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILDATANVE
ALRLLAGLGLQQQLDEGLFSRLRNLHDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLAPGDLSGLFPRLRLLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHIACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRSLTGLIEPVSPSTSLRVGLQRYLQGSSVQLSLRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALA
LLAALAAVGAAYCVRRGRAMAAAQDKGQVGPAGPLEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFPGPGLQSPLHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

CGACGGCGAGGCGGGGTGGCTGAGTCGTGGCAGAGGCGAACAGCTCATGCC
GGTCCGGATAGGGCTGACGCTGCTGTGCGGTGCTGAGCTTGGCTCGCGTCCCG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCAAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTGATTCAAAGAAATCTGA
ATTAGAATCCTTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCTCAAATCAGAAAACAAGGACTATGAAGAGCCAAGAAAAGTA
CGGAAACCAGCTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGAAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGAATGAAAATCCTTAATGAAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATACTGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATGCTTTTATTGGTGTGATTACTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTGAGAAGCTGACTGAGGAAGGCTCTCCAAAGGGACAGACTGCTCTGG
CTTCTGTATGCCCTGGACTGGTGTAAATTCAAGTCAGGCAAAGGCTCTGTATATTACAT
TTGGAGCTCTGGGGCAATCTAATAGCCCACATGGTTGGTAAGTAGACTTTAGTTAGGAAAGGCT
AATAATATTAACATCAGAAGAATTGTGGTTATAGCGGCCACAACCTTTCAGCTTCATGATC
CAGATTTGCTTGTATTAAGACCAAATATTCAAGTGAACCTCCTCAAATTCTGTTAATGGATAT
AACACATGGAATCTACATGAAATGAAAGTTGGAGTCCACAATTTCCTTAAATGATTAG
TTGGCTGATTGCCCTAAAAAGAGAGATCTGATAAATGGCTTTAAATTCTCTGAGTTG
GAATTGTCAGAATCATTTTACATTAGATTATCATAATTAAAATTTCTTAGTTCA
AAATTGTAATGGCTATAGAAAACACATGAAATATTACAAATTGGCAACATG
CCTAAGAATTGTTAAAATCATGGAGTTATTGTGCGAGAACAGACTCCAGAGAGCTCTACTTCTG
TTTTTACTTTCATGATTGGCTGCTTCCATTATTCTGGTCAATTGCTAGTGACACTGT
GCCTGCTCCAGTAGTCTCATTTCCATTGGTCAATTGTTACTTTCTTGCTAATTGG
AAGATTAACTCATTAAATAAAATTATGCTAAGATTAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 18

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLSKTTLTSDESVKDHTTAGRVVAGQIFLDSEESSEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPNENKDYEPPKKVRKPALTAIEGTAHGEPCCHFPFLFLDK
EYDECTS DGDREDGRILWCATTYDYKADEKWGFCTEEEAKRRQMQEAEEMMYQTGMKI LNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLQPNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCA GATTTAAGCCCATTCTGCAGTGGAA TT CATGA ACTAGCAAGAGGACACC ATCTCTT
GTATTATA CAAGAAAGGAGTGTA CCTATCACACACAGGGGGAAAAATGCTCTTTGGGTGCTAGG
CCTCCTAATCCTCTGGTTCTGTGGACTCGTAAAGGAAA ACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTGAAACTTGGCAGCCAGAACTTTGAT
AAAAAGGGATTCA GTAATCGCTGCC TGTGACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTCGTACTGTGCTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTGGGTCTGATCAATAATGCTGGTGT
CCCGGC GTGCTGGCTCCC ACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATGCTCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTGCCCTGCAATCGTGGAGGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGGTTCAATGACAGCCTAACAGACGGACATGAAAGCTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAAACTTGGCAGATCCAGTAAAGGTAAATTGAAAAAAAC
TCGCCATTGGGAGCAGCTGTCCAGACATCAAACAATATGGAGAAGGTTACATTGAAAAAA
AGTCTAGACAAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCCGGTGGTAGAGTG
CATGGACCACGCTAACAAAGTCTCTCCCTAACAGACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTCTGGATA CCTCTGTCTCACATGCCAGCAGCTTGCAAGACTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCAAGGCAGTGTGACTCAGCTAACCAAAATGTCTCCAGGCTATGA
AATTGGCCGATTCAAGAACACATCTCCTTCAACCCATTCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTCTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGTATCCCAGGGT
CCCTGCTCAAGTTCTTGAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTTAGGCTTGCCTGCTTGGTGTGATGTAAGGAAATTGAAAGACTTGCCCATTCAAAATGA
TCTTTACCGTGGCCTGCCCATGCTTATGGCCCTCAGCATTACAGTAACCTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAGATAAGTCACCCAAAAAA
AAAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGKGLWGLINNAGVPGVIAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIKEKKLAIWEQLSPDIKQQYGEKYIEKSLDLKGNKSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKI FWIPLSHMPAALQDFLLLKQKAE LANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGC GGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGGCG
CACTCGCTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTCTTCTGGGGAAAGTAAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATAACAATTGA
CATTCAGAAATATATTCCATGCTATCAGTTTAGCTTTATAATTCTTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATT CAGATCAGATCATGACGTTAGAGAGAGGCTGCTTCACAAAAACTTGCAAGGAGCATT
TTCAAAACCAAGACCTTGTCTGCTATTAAACCCAAGTATAATAACAGAAAGCTGCTCTACTC
ATCGACTGGAACATT CCTTATATAAACCTCAAAAGGACTTTTACAGGGTACCTTAGTGGTT
GCCAATCTGGGCATGTCTGAACAACTGGGTTATAAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTGAGAAGAGATGGATCCTTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAAGCAGTAGATAAACTAGTAAAGGATGTAACACAGATTAAACAGAGAAATTGA
GAAAAGGAGAGGAGCACAGATT CAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
ACATTTTCTTGTCAAGCATTACGGACCTTTTCAAATTCTGAATTCTTCATT CATGTGTT
ATGTCTTAAAAAATAGACATGTTCTAAAGTAGCTGTAACTACAACCACATCTGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATT CCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTGGTGAATATT CACGGCTC
CTACATTTGATCTTTAACCTTACAAGGAGATTTTTATTGCTGATGGTAAAGCCAAAC
ATTCTATTGTTTTACTATGTTGAGCTACTGCA GTAAAGTCTTTGTTACTATGTT CACC
TGTTGCA GTAAACACAGATAACTCTTAGTGCATTACTTCACAAAGTACTTTCAAACATCA
GATGCTTTATTCCAAACCTTTTACCTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
ACACATTCTT AGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACC ATGTCTATTAAAAAATGAAAAGCAAGAATAGCCTTATTTCAAAATATGGAAA
GAAATTATGAAAATTCTGAGTCATTAAAATTCTCCTTAAGTGTAACTTTTGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKNNVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEDGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFFNPSEFLHSCVMSLKNRHSVSKSSCNHHLDVVDMNTL
MVEHTDipeASPASTPQIIKHKALDLDDRWFQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGCCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCAA
GCAGCGCGCAGCGAACCCCCGCCGCCACACCCCTGCGGTCCCGCGGCCCTGCCACCCCTCCCTCCCTCCCC
GCGTCCCCGCCCTGCCGGCCAGTCAGCTTGCGGGTCTCGCTGCCCGCGAAACCCCGAGGTACCCAGCGCGCTCT
GCTTCCCCTGGGGCGCGGCCGCCCTCACGCCCTCCTCTCCCGGCCCTGGCACCGGGACCCTGGCTGA
CGCGAGGCCAGCTACTTTGCCCGCGTCTCTCCCGCTGCGCTCTTCCACCAACTCCAACCTCTTCC
TCCAGCTCCACTCGTAGTCCCGACTCCGCCAGCCCTGGCCCGTGCCTAGCGCCGCTTCCCGTCCGGTCCAAA
GGTGGGAACGCGTCCGCCCGGCCGACCATATGGCACGGTTGGCTTGCCCGCCTCTGCACCCTGGCAGTGC
AGCGCCGCGCTGGCTGCCGAGCTAAGTCGAAAGTTGCTCGGAAGTGCAGCTTACGTGTCAAAGGCTC
AACAAAGAACGATGCCCGCCCTCACAGAGATCAACGGTATCAATTGAAGATCTGCCCCAGGGTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCGTCAAAGTAAGATGATTCAAAGTGTGGTCAAGCGAACAGTGAATCATTG
CAAGCTGCTTGTCTCACGTTACAAGAAGTTGATGAATTCTCAAAGAAACTACTGAAAATGAGAAATCCCTG
AATGATATGTTGTGAACACATATGCCATTATACTGCAAATTCTGAGCTATTAAAGATCTCTCGTAGAGTTG
AAACGTTACTACGTGGTGGAAATGTAACCTGGAAGAAATGCTAAATGACTTCTGGCTCGCTCCGGAGCGGATG
TTCCGCGTGGTGAACCTCCAGTACCACTTACAGATGAGTATCTGGAATGTTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTCCGAGATCTCCCTCGCAAATTGAAGTCCAGGTTACTCGTCTGGTGTAGCAGCCGTACTTCGCTCAAGGC
TTAGCGGTTGCGGGAGATGTCGTGAGCAAGGTCTCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCAGTCCGGGCTCGTGAAGTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTGGCCAACCAAGGGATCTGATTTGAATGAAACAATTCTAGATGCTATGCTGATGGTGGAGAGAGGCTA
GAGGGTCTTCACACATTGAATCGGTATGGATCCCATCGATGTGAAGATTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTCAAGTGTCTCAGAAGGTTTCCAGGGATGTTGAGCCATGTTACAACACTGCTCAAACATCATGAGAGGC
TCCATCTGAAAGTGCCTCAGTGTCTGCTTCAAGACCACATCCCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTTGGACCGACTGGTTACTGATGTCAGGAGAAACTGAAACAGGCCAGAAATTCTGGCTCCCTCCGACCAAC
GTTTGCAACGATGAGAGGATGGCTGAGGAAACGCCAATGAGGATGACTGTTGGAATGGAAAGGCCAAAGCAGGTAC
CTGTTGCAAGTGTGACAGGAAATGGATTAGCCAACCCAGGGCAACACCCAGGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCCTCTGCAAACTGATGGCTCTCGAGTGTGACCGAGAAGATGACATAACATGGAACGACGTG
GACTTCTTGATATCAGTGTGAAAGTAGTGGAGAAGGAAGTGGAGTGGCTGTGAGTATCAGCAGTGCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGAGAGTGGCAATGAGAAAGCCGACAGTGTGCTGGTGTCCGCTGGGCA
CAGGCCTACCTCCTCACTGTCCTGCACTTGTGTTCTGGTATGCAAGAGAGAGTGGAGATATTCTCAAACACTGAG
AAAAAGTGTCAAAAGTAAAGGACCAAGTTACCTTACATGGCTCTGGTAAAGAAGTGTGACTTGTGTTCTCATTGAGTTGG
TGGACAACAAATGTACAGTTTACTATGTCGACTGGTAAAGAAGTGTGACTTGTGTTCTCATTGAGTTGG
AGGAAAGGGACTGTGATTGAGTTGGTCTGCTCCCCAAACCATGTTAACAGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCAATTGTGATTCTACTCTATTATTTGTTGATGTTTTCTCATTGTTGTGG
TTTTTTCAACTGTGATCTCGCCTTGTCTTACAAGCAAACCAAGGGTCCCTTGTGGCACGTAACATGTACGGTATT
TCTGAAATTTAAAGTGTACAGAAGCAGGTTTATTATGTTATCTTAAAGAAAAAGCCAAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKCSEVRRLYVSKGFKNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSNDMFVKTYGH
LYMQNSELFKDLFVELKRYYVVGNVNLEEMLNDFWARLLERMFRFLVNSQYHFTDEYLECVSKYTE
QLKPFGDVPRKLKLQVTRAFVAARTFAQGLAVAGDVSKVSVNPTAQCTHALLKMIYCSHCRL
VTVKPCNYCSNIMRGCLANQGDLDPEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFOQCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSLPSNVNCNDERMAAGNGNEEDDCWNGKGKSRYLFAVTGNGLANQGNNPENVQDTS
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEKGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCTGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCCATTATATTCTTCAGCAACT
TACAGCTGCACCGACAGTTGCGGAAAGTTCTAATCTTCCCTCTCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTCAAAGATTGGTCTGAG
AGCCCCGAGAAGAAAATTATGACAGTGTCTGGGCTGCCAAGAAGCAGTGCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAACAAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGCTCTGCCTTGTAGGAGCTCTG
AGCGCCCACTCTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCCAGACACTC
TTCTTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAGCA
TGTTTTCAAGATCATTGTTGCTCTCTAGTGTCTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTAACCGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCC
AGCTAGTGTCAATTAAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTT
AAATGTCAAAAAAAAAAAAAAA

FIGURE 26

MKVLISSLLLLPLMLMSMVSSSLNPVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKKQCPDCDFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCACCTCTGCAGCTGCTGGTGTGCTTCTTACCCCTGCC
CCTGCACCTCATGGCTCTGCTGGCTGCTGGCAGCCCTGTGCAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGAAGATGGAGAGCAAGAAACGGAGCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGAAAGTGGCCCTACTGGAGCTGGCTGCCAACCGAGC
CAACTTCAGTTCTACCCACCAGGGCTGCAGGGTACCTGCCTAGACCCAATCCCCACTTGAGA
AGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCAATATGAGCGGTTGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGTGCACTCTGGTGTG
CTCTGTGCAGAGCCAAGGAAGGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCCGGAGGTGTG
TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGCCTTCATGTGGCAGCAAGTTTC
GAGCCCACCTGGAAACACATTGGGGATGGCTGCCTCACCAAGAGAGACCTGGAAGGATCTTGA
GAACGCCAGTTCTCGAAATCCAATGGAACGACAGCCCCCTCCCTGAAGTGGTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGCAAACAATCTTCCAAGCTCCAAGGCACACTTTGCTCC
TTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGCTATCTATCTTCCACTGAGAGGGACTA
GCAAGATGAGAGAACACATTGATGTACCACTACTAGTCCCTCTCCCCAACCTGCCAGGGC
AATCTCAACTCAATCCGCCCTCGACAGTAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGATCCTCACTGCAAGTTCTGGACTAGTCTCCCAACGTTGCTC
CCAATGTTGTCCTTCCCTCGTTCCATGGTAAAGCTCCTCTCGCTTCCCTGAGGCTACAC
CCATGCGTCTCTAGGAACCTGGTCACAAAGTCATGGTGCCTGCATCCCTGCCAACCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTGAGCTGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVVAPGEDM
RQLADGSMVVVCTLVLCVQSPRKVLQEVRRLRPGGVLFFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFPSSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTAAAGGCTCATGCTGGAGTGGGACTGGTCGGGCCAGAAAGTCTCTCTG
CCACTGACGCCCATCAGGGATTGGGCCTCTTCCCCCTCCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGAGAAAGTGGGGATGGTA
AGAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGTGGCTAGGGGGCTGCCTTATTTAAA
GTGGTTGTTATGATTCTATACTAATTACAAAGATATTAAGGCCCTGTTCATTAAGAAATT
GTTCCCTTCCCTGTGTTCAATGTTGTAAGATTGTTCTGTGTAATATGTCTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTC
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTCAGATTCCCTCAACTATAACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTGATAAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGAATCATGTCGG
GAAGAGATAACAATCCTGGCCTGTGTATCCTCGCATTAGCCTGTCTTGGCCATGATGTTACC
TTCAGATTCATCACCAACCCCTGGTTCACATTTCATTGGTTATTTGGGATTGTTGTT
TGTCTGCGGTGTTATGGTGGCTGTATTATGACTATAACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGAAAATATGAAGTCCGTGCTGGGTTGCTATCGTATCCACAGGCATACGGCAGTG
CTGCTCGTCTGATTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGTCCCTGGTGGCTGCTGAGCCTGGAAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGGAAATATAAGCCCCCTTCGGCATTGGTACATGTGGTGTGTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTGCGTGCCAGCAAATGACTATAGCTGGGCAG
TGGTTACTGTTATTCAACAGAACTAAAAATGATCCTCTGATCATCCCATTCTCGTCTCTC
TCCATTCTCTTCTACCATCAAGGAACCGTTGAAAGGGTCACTTTAATCTCTGTTGAG
GATTCCGAGAATCATTGTCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTCGATGCTGCTACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACCGAAATGCATAACTACAACGTCAATTAAATGGGACAGATTCTGTACATCAGCAAAGATGC
ATTCAAATCTTGTCCAAGAACACTCAAGTCACTTACATCTATTAACTGCTTGGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGTGGTGTGTTCACTGTTGGAGGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGTGGCAGTCCCTCTGTTATTGGTAGCTTTGCCTACTTAGTAGC
CCATAGTTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCTGTGTTGCTGTTGATC
TGGAAACAAATGATGGATCGTCAGAAAAGCCACTTATGGATCAAGAATTCTGAGTTCGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATAGAACCCATTAGGTATCTGTACCTGGAAAACATT
TCCTTCTAAGAGCCATTACAGAAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATT
TTAAAAGACCTAATAAACCTATTCTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSIAMMFTFRFITTLLVHIFISLVLGILFVCGVWWLYYDYNLDSIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDDPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMMQMNALKEQQHG
ALSRYLFRCYCFCFWCLDKYLLHLNQNAYTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFCGD
FIIFLGKVLVVCFVFGLMAFNYNRADFQWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTGGACCTCTCCCTGTTCTCCTAGA
ATAATTGTATGGGATTGTGATGCAGGAAGCTAACGGAAAAAGAATATTCAATTCTGTGTGGT
GAAAATTGGGAAAGGGCTAACGGCTAACGGAAAAAGAATATTCAATTCTGTGTGGT
TGTTGTCTCACTATGAAGGCATCTGTATTGAATGTTCTGTTGCTGGTGAETGGAGTAC
ATTCAAACAAAGAACGGCAAAGAAGATTAAAGGCCAACGTTACTGTGCCTCAGATCAACTGC
GATGTCAAAGCGGAAAGATCATCGATCCTGAGTTCTGTAAATGTCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTATGCATCCTACTCCAGTGTGTGGCGTGC
TACACAGTGGTGTGCTGATAATTCAAGGAGGGAAAATACTTGTTCGGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCAATCGTTATCCCTACCACGATGGAGAGAACTCCTT
TATCGTCTTAGAAAGTAACCCAAAAGGGTGTAACTTACCCATCAGCTTACATACTCATCAT
CGAAAAGTCCAGTCCCAGCAGGACGACCAAAAGCTATCAGAGGCCACCTATTCCAGGG
ACAACAGCAGCCGGTCACTCTGATGCAGCTCTGGCTGTCACTGTAGCTGGCCACCCAC
CACCTGCCAAGGCCATCCCTCTGCTGCTCTACACCAGCATCCCCAGACCAATCAGTGG
GCCACAGGAGGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAAGGCC
AGAGCTGATCCAGGTATCCAAGGCAAGATCCTCAGGAGCTGCCCTCAGAAACCTGTTGGAGC
GGATGTCAGCCTGGGACTTGTCCAAAAGAAGAATTGAGCACACAGTCTTGGAGCCAGTATCCC
TGGGAGATCCAAACTGCAAATTGACTTGTGTTTTAATTGATGGGAGCACCAGCATTGGCAA
CGGCGATTCCGAATCCAGAAGCAGCTCTGGCTGATGTTGCCAACGCTTGAATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCAGTATGGAGACAACCTGCTACTCACTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTCTAAT
GTAGGTCGGGCCATCTCTTGTGACCAAGAACCTTCTTCCAAAGCCAATGGAAACAGAACGG
GGCTCCAATGTGGTGGTGGTGTGGTGGATGGCTGGCCACGGACAAAGTGGAGGAGGTTCAA
GACTTGCAGAGAGTCAGGAATCAACATTCTCATACCATTGAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGCCCAACTTGCACAAACAAGGGCGTGTGAGAACAAACGGCTTACTC
GCTCCACGTGCAGAGCTGGTTGGCCTCCACAAGACCTGCTGAGCCTCTGGTGAAGCGGGTCTGCG
ACACTGACCGCCTGCCCTGCAGCAAGACCTGCTGAACTCGGCTGACATTGGCTTCATCGAC
GGCTCCACGTGTGGGAGCGGCCACTTCCAGCGTGTGACCAACCTCACCAA
AGAGTTGAGATTCCGACACGGGACAGCAGCAACTTCCGACAGCAGCAAGCCTGAGTACACCTACGAACAGCGGC
TGGAGTTGGGTTGACAAAGTACAGCAGCAAGCAGCAGCAACGGGCTGCCATCAACTCGCCCTGGAGCAGCTTCAAGAA
TACTGGAGTGGTGGCACCAGCACGGGGCTGCCATCAACTCGCCCTGGAGCAGCTTCAAGAA
GTCCAAGCCAAACAAGGAAAGTTAATGATCCTCATCACCAGCGGAGGTCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCCATCTGAAGGGAGTGTACACCTATGCGATAGGCGTTGCCTGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCACTCCTTCTTGTGGACGA
GTTTGACAAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTGTACAGAGTTCAACTCAC
AGCCTCGGAAC**TG**AATTCAAGCAGCAGGACCCAGCAAGTGTGCTTACTAACTGACGTGTT
GGACCCACCCACCGCTTAATGGGGACGCACGGTGATCAAGTCTGGCAGGGCATGGAGAAC
AAATGTCTGTTATTATTCTTGCATCATGCTTTCATATTCAAACATTGGAGTTACAAAGA
TGATCACAAACGTATAGAATGAGCAAAGGCTACATCATGTTGAGGGTGTGGAGATTACAT
TTTGACAATTGTTCAAATAATGTTGCAATACAGTGCAGCCCTACGACAGGCTTACGTAG
AGCTTTGTGAGATTTTAAGTGTATTCTGATTGAACTCTGTAACCCCTCAGCAAGTTCA
TTTGTCATGACAATGTAGGAATTGCTGAATTAAATGTTAGAAGGATGAAAATAA
AAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 34

MRTVVVLMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLNDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTVPSALTYSSSKSPAAQAGETTKAYQRPIPGTTAQPVTLMQLLAVTVAVA
PTTTLPRPSPSAASTTSIPRPFQSVGHRSQEMDLWSTATYTSSQNRPRADEGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPMFANKAVCRTNG
FYSLHVQSWFGLHKTLQPLVKRVCDDRLACSKTCLNSADIGFVIDGSSVGTGNFRTVLQFVTN
LTKEFEISDTDRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
FKKSXPKRKLMILITDGRSYDDVRIPAMAHLKGVITYAIGVAWAQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCCTTAGGAGGTGGCTGCCTGTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGCTTTTCTGTTTCAGAGTAGTTACAACAGATCTGAGTGTTAATTAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTCTCATCTGGAATTCCACAGTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTCCCGCTGGCTGCTCTACACGTGGTGCTCTCCGACTACTCACCCGAGTGT
AAGAACCTCGGCTCGCGTCTTGAGCTGCTGTGGATGGCCTCGGCTCTGGACTGTCTCCGAGTA
GGATGTCACTGAGATCCCTAAATGGAGCCTCCTGCTGTCACTCCTGAGTTCTTGATGTGGTAC
CTCAGCCTCCCCACTACAATGTGATAGAACCGGTAACTGGATGTACTCTATGAGTATGAGCCGATT
CAGACAAGACTTCACCCACACTTCAGAGAGCATTCAAACAGTCTCATCAAATCCATTCTGGTCATT
TGGTGACCTCCACCCTCAGATGTGAAAGCCAGGCCATTAGAGTTACTTGGGTGAAAAAAAGTCT
TGGTGGGATATGAGGTTCTTACATTTCTTATTAGCCAAGAGGCTGAAAGGAAGACAAATGTTGG
ATTGTCCTTAGAGGATGAAACACCTTTATGGTACATAATCCGACAAGATTTTAGACACATATAATA
ACCTGACCTTGAAACCATTATGCCATTAGGTGGTAACGTGAGTTTGCCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTCATCAAACTGGCAATTAGTGAAGTATCTTAAACCTAAACCACTCAGA
GAAGTTTCACAGGTATCCTCTAAATTGATAATTATTCTATAGAGGATTACCAAAAAACCCATATT
CTTACCAAGGAGTATCCTTCAAGGTGTTCCCTCCACTGCAGTGGGTTGGTTATATAATGTCCAGAGAT
TTGGTCCAAGGATCTATGAAATGATGGTCACGTAAACCCATCAAGTTGAAGATGTTATGTCGGAT
CTGTTGAATTATAAAAGTGAACATTCAATTCCAGAACACAAATTCTTCTTATAGAATCC
ATTGGATGTCCTGCAACTGAGACGTGTATTGCAGCCCATTGGCTTTCTTCAAGGAGATCATCATT
TGGCAGGTATGCTAAAGAACACCACATGCCATTAT**TA**ACTTCACATTCTACAAAAAGCCTAGAAGGACAG
GATACCTGTGAAAGTGTAAATAAAAGTAGGTACTGTGAAAATTCAATGGGAGGTCACTGTGCTGGCTT
ACACTGAAACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG
CCCTCAAAAGATGATATGTGGAGGAATTAAATATAAGGAATTGGAGGTTTGCTAAAGAAATTAAAGG
ACCAAAACAATTGGACATGTCAATTGTAGACTAGAATTCTAAAAGGTGTTACTGAGTTATAAGCTCA
CTAGGCTGTAAAAACAAACATGTAGAGTTATTATGAAACAATGTAGTCATTGAAGGTTTGTA
TATCTTATGTGGATTACCAATTAAAAATATGTAGTTCTGTGCAAAAAACTCTTCACTGAAGTTATA
CTGAACAAAATTTCACCTGTGTTTGGTCATTATAAAAGTACTTCAGATGTGCACTGTTACAGTTATT
ATTATTAAAAATTACTCAACTTGTGTTTAAATGTTTGACGATTCAACAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTCCAGTTACTTAACAGTACAGTTATTATTGATAACATCACTCCA
TTAATGTAAAGTCAGGTCAATTGGCATATCAGTAATCTTGGACTTTGTTAAATATTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLLKWSLLLSSLFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHNSNC SHQN PFLV I LVTSHPSDV KARQ A I RV TWGEKK SWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDII RQDFLDTYNNL TLKTI MAF RWVTE FCPNAKYVMKT DTDV FINTGNLV KYLL
NL NHSEK FFTGYPLIDNYSYRGFYQKTHIS YQEYPFKVFPPYCSGLGYI M S RD L V P R I Y E M M G H V
KPIKFEDVYVGICLNLLKVNIHIPE D TNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAAGGAATGGAGCTGGTTGCTGGACGCAGTTGGGCTCACTTTCTCAGTCCTTCATC
TCGTCCCTGCCAAGAGAGTACACAGTCATTAAATGAAGCCTGCCCTGGAGCAGACTGGAATATCATGTCGGAGTGCTGTG
AATATGATCAGATTGAGTCCGTCTGCCCGGAAAGAGGAAAGCTGCTGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCCTGCCGTATCCACCCAGGTTGATCATTCTTGAAGAAACTGCAAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTTGGATGACTCTATGTGAAGGGGTTCTACTGTGAGGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTGTTGAGAAGCTATCCCCTAAATGTCAGTGTGAATGGACCATTGATGC
TAAACCTGGGTTGTCATCCAACTAAGATTGTCATGTTGAGTCTGGAGTTGACTACATGTCAGTATGACTATGTTGAG
GTTCGTGAATGGAGACAACCGCGATGCCAGATCATCAAGCGTGTGTCAGGCAACGAGCGGCCAGCTCTATCCAGAGCATAG
GATCCTCACTCCACGTCCTTCCACTCCGATGGCTCAAGAATTGACGGTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCTCATCCCCGTTCCATGACGGCACGTCGCTTGCACAGGCTGGATCTTACAAGTGTGCTGCTGGCAGGC
TATACTGGGAGCGCTGTGAAAATCTCCCTGAAGAAAGAAACTGCTCAGACCCCTGGGGCCAGTCATGGGTACAGAAAA
TAACAGGGGGCCCTGGGTTATCACGGACGCCATGCTAAAGGGCACCGTGGTGTCTTCTTTGTAACAACCTCTATGT
TCTTAGTGGCAATGAGAAAAGAACTTGGCAGCAGAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCTGCCGA
GAACCAAAGATTCAGACCTGGTGAAGAGAGTCTTCGATGCCAGGTTGACTCAAGGGAGACACCATTACACAGCTAT
ACTCAGGGCCTTCAGCAAGCAGAAACTGCAAGTGCCTTACCAAGAAGCCAGCCCTTGGAGATCTGCCATGG
ATACCAACATCTGCATACCCAGCTCAGTATGAGTCATCTCACCTCTACCGCCGCTGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGCACCCTGCTGCATCCCTATCGGGAAAATTGAGAACATCACTGCTCCAAAAGA
CCCAAGGGTTGCGCTGGCGTGGCAGGCCATCTACAGGAGGACCAGCGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCTAGTCTGCAGCGGTGCCCTGGTGAATGAGGCGACTGTGGTGGTGCCTGGCAGTCAGTGTACTGACCTGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTGGGGAAATTCTACCGGGATGATGACCGGGATGAGAACACATCC
AGAGCCTACAGATTCTGCTATCATTGCTCACCTGCTGATGCTGACATGCCATCTGAAGCT
CCTAGACAAGGCCGTATCAGCACCCGAGTCCAGCCATCTGCCTGCTGCCAGTCGGGATCTCAGCACTCCCTCAGGAG
TCCCCACATCACTGTGCTGGCTGGAATGTCCTGCCAGAGCTGAGGAGGCCCCTGGCTCAAGAACACACTGCGCTCTGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGTGGGAACCCACTGCCCTCTGATATCTGCACTGCAAGAGACAGGGAGCATGCCAGTCGGTGTGCTTCCCGGA
CGAGCATCTCTGAGCCACGCTGGCATCTGATGGGACTGGTCACTGGAGCATGATAAAACATGCAAGCCACAGGCTCTCCA
CTGCCCTCACCAAGGTGCTGCCCTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACACTCTGAGAAG
TGTTTCTGTATATCCGCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAAGTGTGATTGGCTGTGAATTGGCT
GTGCCAGGGCTCTGACTTCAGGGACAAAACACTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGAGATGCCAGGGCTGCAAGAAGTAAGTTCTCAAAGAACCCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGTCTCCCAACTTCAGTTACGAATGCCATCAGCTTGACAGGGAGATCTGGGCTCATGAG
GCCCTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCGTGGGACAGCCCAGGGAGCAGAGCTGGGATGTGGTGCATGCC
TGTGTACATGGCCACAGTACAGTCTGGCTTTCTCCCATCTCTGTACACATTAAATAAGGGTTGGCTTCT
GAAACTACAA
AA

FIGURE 38

MELGCWTQLGLTFLQLLLSSLPREYTVINEACPGAEWNIMCRECCYEVDQIECVCPGKREVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQQNGEWSGKQPICIKACREP KISDLVRRRVLPMQVSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPGMYQHLHTQLQYECISPFYRLGSSRTCLRTGKWSGRAPSCIPI CGKIEINITAP
KTQGLRWPWQAAIYRRRTSGVHD GSLHKGAFLVC SGALVNERTVVVAAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDREDETIQSLQISAIILHPNYDPILLADIAILKLLDKARI STRVQ PICLAASR
DLSTS FQESHITVAGWNVLADVRSPGFNDTLRG VSVVDSLLCEEQHEDHGIPV SVTDNMFCA
SWEPTAPS DICTAETGGIAAVSF PGRASPEPRWH MGLV SWSYDK TCSH RLSTAFTK VLP FKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCCTACATCCTCATCTGAGAACATCAGAGAGCATAATCTTACGGGCCGTGATTATTAACGTGGCTTAATC
TGAAGGTTCTCACTCAAATTCTTGTACTGATCTACTGATGTGGGGCATGGCAAGGTTGCTTAAGGACCTGGCTGG
TTTGGGCCCTTGCTAGCTGACAGAAGGGCCAGGAGAAATGCAGCACACTGCTCGAGAATGAAGGCGCTTCTGTTGC
TGGCTTGCCCTGGCTCAGTCCTGCTAACTACATTGACAATGTGGCAACCTGCACCTCCTGTATTCAAGAACTCTGTA
AAGGTGCCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCCTCACAAAGATGGCTGTCCAGACGGCTGTGCCAGCC
TCACAGCCACGGCTCCCTCCCCAGAGGTTCTGCAGCTGCCACCATCTCTTAATGACAGACGAGCCTGGCTAGACA
ACCCCTGCCCTACGTGTCTCGGCAGAGGACGGGCAGGGCAGCAATCAGCCAGTGGACTCTGGCGAGCAACCGAACTA
GGGCACGGCCCTTGAGAGATCCACTATTAGAACAGATCATTAAAAAAATAATCAGGCTTGAGTGTCTCGAA
GGACAAAGAGCGGGAGTCAGTTGCCAACCATGCCGACAGGGCAGGGAAAATTCTGAAAACACCACTGCCCTGAAG
TCTTTCAAAGGTGTACCACCTGATTCAGATGGTAAATTACCGCATCAAGATCAATCGAGTAGATCCCAGTGAAA
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCCACTGGTCCATATCATTATCCAACACATTATCGTATGGGG
TGATGCCAGAGACGGCCGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACACTACGCTGTGCGTCTCTGCCGAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTCCGCA
GCAGGAACAATGGACAGGCCCCGGATGCCCTACAGACCCCGAGATGACAGCTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGGACGAGCTTGGAAATAAAACTCGTGCCTAGGGATGACGCTGGGTTTCATCTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCA
GCCCACAAAAGTGCGGCTCATCTGATTCAGGCCAGTGAAAGACGTGTTCACCTCGTGTGCCGCCAGGTTCGGCAGC
GGAGCCCTGACATCTTCAGGAAGCCGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCAGGGAGAGCAACA
CTCCCAAGCCCCCCTACCTACAAATTACTGTCACTGAGAACGGTGGTAAATATCCAAAAGACCCGGTAATCTCTCG
GCATGACCGTGCAGGGGGAGCATCACATAGAGAATGGGATTTGCTATCTATGTCACTCAGTGTGAGCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTGTTGAATGTGGATGGGATGGCTGAACACTGACAGAGGTCA
GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTGGAAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAAACATGGCCCCACCCAGTGA
GGTCACTGTGGCTGGAAATTACACCGTGCTGTATAACTGTAAAGATATTGTATTACGAAGAACACAGCTGGAGTC
TGGGTTCTGCATTGTAGGAGGTTATGAGAACATGAAACAAACCTTTTCTACAAATCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTATTCCTCTGCTGTCAATGGTAAAGTACATCAGGAATGA
TACATGCTTGTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTCTGGCTGGCACTT
TTTATAGAATCAATGATGGTCAGAGGAAACAGAAAATCACAATAGGCTAAAGTGAAGTAAACACTATATTATC
TTGTCAGTTTATATTAAAGAAAATACATTGTAAGGAAAGTGTCACTGAGGAAAGTATGATCATCTAAAGGCCAGTT
ACACCTCAGAAAATATGATCCAAAAAAATTAAACACTAGTTTTTCAGTGTGGAGGATTCTCATTACTCTAC
AACATTGTTATATTCTATTCAATAAAAGCCCTAAACAAACTAAAATGATTGATTGTATACCCACTGAATT
CAAGCTGATTAAATTTAAAGGTATATGCTGAAGTCTGCCAGGGTACATTGGCCATTAAATTACAGCT
AAAATTTTAAATGCTGAGAACGTTGCTTCATCAAACAGAATAATTTTCAAGTAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDCASLTATAPS
PEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGRSNRTRARPFERSTIRSRFFKKINR
ALSVLRRTKSGAVANHADQGRENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFI FNVLDDGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSRQVRQSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVNNIQKDGPESLGMVTAGGASHREWDLPIYVISVEPGGVISRDRG
IKTGDILLNVDGVELTEVSERSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSHNMMAPP
SDWSPSPWVMWLELPRLYNCKDIVLRRNTAGSLGFCIVGGYEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLRLLKELKGRITLTIVSWPGTEI

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAAGGCATTGTATCTTCAGTTGTCATCAAGTTGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCTGAGTCAGTGAAGCAGAGAGATAGATATTATTACAGTAATAAAAACATGGGCTTCAACCTGACT
TTCCACCTTCCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTCCTGCACAGTGGTTGGTGGC
CACCAAGTAACACTTCGTGGGTGCCATTCAAGAGATTCTAAAGCAAAGGAGTTCATGGCTAATTCC
ATAAGACCTCATTTGGGAAGGGAAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACCT
GACAACGTCCCTCTGTGTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTCAAACCAAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCAAAGTGTCCAGAGGCCGTATGCCCTCAGGAATGTAAG
CTTACAGAGGGTCGCCATCCTCGTCCCCACCGGAACAGAGAGAAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTTCCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCACTCCACCAGGCTGAAGG
TAAAAAGTTAATCGAGCAAACCTCTGAATGTGGCTATCTAGAACCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTAACTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTGGCAGGAACAGCACTGGTACAGGTTACGTTACAGTGGATATTGG
GGGTGTACTGCCCTAACGAGAGCAGTTCAAGGTGAATGGATTCTCTAACAACTACTGGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTGAGCTCCAAGGAATGAAAATTCCCGCCCCCTG
CCTGAAGTGGTAAATATAACATGGTCTTCCACACTAGAGACAAAGGAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTACAGACTGGGAGAACAGATGGTTGAGTAGTTGTTCTTATAAAAT
TAGTATCTGTGGAACACAATCCTTATATATCACACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTGGTGTGTTGGAAAGAACACTGATTCTTGTGCAATAATTGGCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTCCTTTGTATTTCT
TAGCAGAGCTCCTGGTGTGAGTAGTATAAAACAGTTGTAACAAGACAGCTTCTAGTCATTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTGAAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTGAGGAGATTATTTAAATTGAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCAGAGAACCCAGAGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAATACTGTTATTCAATTATCCTGTAACATCTGTGAAGTGGTGGTGTCAAGGT
GAGAAGGCGTCCACAAAAGAGGGAGAAAAGGCAGCAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAGGCAGCAGTAGCTGAGCTGGTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCCTCAGTGATGCCACCAGAGAATACATTCTTATTAGT
TTTAAAGAGTTTGTAAGGATGATTGACAAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAAAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLIGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFIHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVITALSREQFFKVNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGTGGACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTCCCACAAACAG
ACGGGACAACTTGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCAT
GTTCCAGAGGCAGAGGAGGCAGACACCCACTTCCCCATCTGCATTTCCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCCCCTCCC
TTCCTTATTATTCTGCTGCCAGAACATAGGTCTTGAATAAATGGCTGGTTCTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMMPMFQRRRRDTHFPI
CIFCCGCCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTTGTACCACAGCCAGAAGGGGGACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATAACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGACACCTGTCAAACCCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGTATGACCTTCATCTGCCAGGAAGCTCTGTGAAGGTGCTGATGACCCAGATTCCCTCCATG
TCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGATGACCCAGATTCCCTCCATG
GTCCTCCTGTGTCTCCTGTTGGTGCCTCCTGCTCAGTCTTTGACTGGGCTATTCTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGGTGGACATTGTCGGG
AAACTCCTAACATGCCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAAATCCTAAAGGAAGATCCAGCAAATACGGTTACTCCACTGTGGAAATACGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCCTAAGTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSGAVTFPLSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLFLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTCTACTGCTGTTAGGAGTAGTTCAATGCGATACCCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCTCAAAAACCCATCTCTGCTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAAACAGAACTGGAATGTTCTTCATCATTTTCAGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTAAAAGGTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATGTAATTTTCATTGAAAAACATCAGTGACATTCACCAGAATCCTCAA
CTTGCAGTGGTTTCAATGACTCTTGTGACCTCCTACTGGTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTAGGTCTATTGCTTGGAAATTCTGGAGGTCTGTTGGCT
CAGTCAGATAGTCATCGGTTCCCTGGCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGGAATAAAATGTAAGTATCAGTAGTTGAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCGCCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTTGGCTGCAGACCCCCGCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTGAACTGTGA
CATGGAGAGAGTGACCCTGGCCCTCTCCTACTGGCAGGCCTGACTGCCTTGAAGCCAATGACC
CATTTGCAATAAGACGATCCCTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCAGGGCTCCTGGCATTGCTGGATCGCGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACGACAGTCCTGTACCTGAGAAGGCCATCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCAGCACC
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCCT
TTCTGATCAGGAGGCTTCTTATGAATTAAACTGCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALLEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGCACAGAGGAGGCCACAGAGACCCAGAGCAAGGGCGGCAAGG
AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGATGAAGTCCAG
GGGCCCCCTGGCCTGCCCTCTGCTGGCCCTGCGCTGGCAGTGGGAGGCTGGCCCTGAGAG
CGGAGAGGAAAGCAGTGGACAAATATTGGGAGGCCCTGGACATGCCCTGGAGACGCCCTGA
GCGAAGGGGTGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGGCCAAGGGACCAGAGAACAGCAGTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGC
AGCAGATGCTTGGCAACAGGGTCGGGAAGCAGCCATGCTCTGGAAACACTGGCAGGAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCGCGGCTCTGGCAGGGG
GTGCCTGGCCACAGTGGTGGCTGGAAACTCTGGAGGCCATGGCATCTTGGCTCAAGGTGG
CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCAGCTTGGATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGG
GGGCCACAAAACCTTGGGACCAACACTCAGGGAGCTGGCCAGCAGCAGTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGGCTCACAGTCGGCAGCAGTGGCAGTGGCAGCAATGGTACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGTGGCGCAGCAG
TGGCGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
CCTCCTGGGATCCAGCACCGGCTCCTCCGGCAACCACGGTGGGAGCGGGAGGAAATGG
CATAAACCCGGGTGTGAAAGCCAGGGAAATGAAGCCCGGGAGGGGAATCTGGGATTCAAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATGCCCTCCTT
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCGAGCTGGGAGTGGAGGAGGTGACCTGTT
GGTGGAGTCATACTGTGAACTCTGAGACGTCTGGATGTTAACCTTGACACTTCTGGAA
GAATTTAAATCCAAGCTGGTTCATCAACTGGGATGCCATAAACAGGACCAAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACAGATTGGATGGAGCCCCACACTCCCTCCTAA
AACACCACCCCTCTCATCACTAACTCAGCCCTGCCCTGAAATAACCTTAGCTGCCCAACAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEVGKAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVP
GHSGAWETSGGHGIFGSQGGLGGQGQNPGLGTPWVHYPGNSAGSFGMNPQGAPWGQGGNGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGSSNSGGSGSQGSSGSGSNGDNNNGSSSGS
SSGSSSSGSSGGSSGGSSGSSGSSGNSGGSRGDGSESSWGSSSTGSSSGNHGGSGGGNGHKGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLGGSGDNYRGQGSSWGSGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGACAAGCTGCTCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGGTGTGGCTCCCTGGCT
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTCTATAACAACTGCCGCCGCTCCAGTGTTCC
CACAGCCCCAAAACGGAACCTGGTTGGGTACCTGGGCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCACCTATTCCCAGGGCTTACGGTATGGCTGGTCCCAT
CATCCCCTCATCGTTTATGCCACCTGACACCATCCGTCTATCACCAATGCCCTAGCTGCCA
TTGCACCCAAGGATAATCTCTTCATCAGGTCCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCCTGGATGCTGACGCCGCCCTCCATTCAACATCCT
GAAGTCCTATATAACGATCTTCAACAAGAGTGCAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCGGACATGTTGAGCACATCAGCCTCATGACCTGGACAGTCTA
CAGAAATGCATCTTCAGCTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTGTAGAGAAAAGAACGCCAGCATATCCTCAGCACATGGACTTCTGT
ATTACCTCTCCCAGTACGGGCGGCCTCCACAGGGCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTCATCCGGGAGCGCGTGCACCCCTCCCCACTCAGGGTATTGATGATTTCAAAGACAA
AGCCAAGTCCAAGACTTGGATTTCATTGATGTGCTGCTGAGCAAGGATGAAGATGGAAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGTCTGTACAACCTTGCAGGCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCC
AGCTGCCCTCCTGACCATGTGCGTGAAGGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATC
TCCCAGTGCACCCAGGACATTGTTCTCCAGATGGCCAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTATTCCCTTCTCCGCA
GGGCCAGGAACGTGACATGGGAGGCAGTCGGCATGGGGAGATGAAAGTGGCTGGCGTTGAT
GCTGCTGCACTTCCGGTCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTTGCAGTGACTTCTGAC
CCATCCACCTGTTTTGCAGATTGTCATGAATAAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWLGRLPVAMSPWLLLLVVGSWILLARILAWTYAFYNNCRRLQCFPQPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPTIRSI TNASAAIAPKDNLFIRFLKP
WLGEGLLSGGDKWSRHRMLTPAFHFNIILKSYITIFNKSANIMLDKWQHLAGSSRLDMFEHI
SILMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRC
RLVHDFTDAVIREERRTLPTQGIDDFFKDKAKSKTLDFIGVLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLRHPEYQERCRCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPPEVYDPFRFPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGCTGTTGTGTTGCACITACCCTGTGTTGCCT
TTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTGCAGTCTTGCGATTGACGTGG
TACAGCCTTCCTTCATAACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTCTTGC
ATAATTCATGCCAGTTTATGAAGCTTGAAGGCACTATGGACAGAACAGCTGGTGGACAGTTT
GTAACATCTCGAACACCTCTGCTTACAGACATGTGCCTTTATCTTGCAGCAATGTGTTGCTT
GTGATTGAAACATTGAGGGTACTTTGGAAGCAACAATACATTCTGAACCTGAATGTCAAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAAATCTCCTCATGTACCTGTTCC
TCTGGATGTTGTCCCCACTGAATTCCCATGAATAACAAACCTATTCAACAGCAACAGCAAAAAAAA
AAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCGGAATCGGCTCGAGGGGCAGTGGAGCACCCAGCAGGCCAACATGCTCTGTCTGCCCTG
TACGTGCCGGTCATCGGGAAAGCCCAGACCGAGTTCCAGTACTTGTAGTCGAAGGGCTCCCTGCCAGCTGAAGTCC
ATTTCAGCTCAGTGTCTCATCCCCCTCCAGGAATTCTCACCTACCGCACTGGAAAGCAGAAAATTGACAAGCT
GGAGATAAGGACCTTGATGGGAGCTAGACTTGAAGAATTGTCCATTATCTCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTTGGACAAAAAGAATGATGGACGCATTGACGCCAGGAGATCATGAGTCCCTGCCAGCTTG
GGACTCAAGATACTGAACAGCAGGAGAGAAAATTCTCAAGAGCATGGATAAAACGGCACGATGACCATCGACTGG
AACAGTGGAGAGACTCACACCTCTCCACCCGTGGAAAATCCTCCAGGAGATCATCTACTGGAAGCATTCCACG
ATCTTGATGTGGTGAAGAATCTAACGGTCCCGATGAGTTCACAGTGAGGAGAGGGCAGACGGGATGTGGGAGA
CACCTGGGGAGGG
CAGGTCAATGCCCTCCCGAGCAACAAACATGGGCATCTGGGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGG
TCACTCTGGGGGGCAATGGCATACAGTCATGGCCCAAAATTCACTGGCCATGAGCAG
ATCAAGGGCTTGTGAGTGGAGGAGACTCTGAGGATTACAGAGGGCTTGTGGCAGGGTCCITGGCAGGGCC
ATCGCCAGACAGCATCTACCAATGGAGTCTGAAGACCCGGATGGCGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGG
ATGCTGGACTGCCAGGG
GGCATCATCCCCATGCCGGCATGCCCTTGAGTCTACGAGACGCTCAAGAATGCCCTGGCTGCAGCACTATGAGTG
AACAGCGGG
CCCCCTGGGGCTAGTCAGGACCCGG
TTCAACATATCTCGGGGACGG
GCTGTGAGCATCAGCTACGTGGCTACGAGAACCTGAAGATCACCTGGGGCTGCACTGGGGGGGGGGGGGGGGGG
CGCCCGGAGTGGACTCGCTGATCTGGGGCCGAGCTGGGGGTGTGCAAGGATCTCATTCTGTGAATGTGCCAACACT
AAGCTGCTCGAGCCAAGCTGTGAAACCCCTAGACCCACCCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
GTCTGCTGACCCAGGACCCCTCTGTGGTTCAGCGAAGACACGGGATCTTAGGGGCTCAGGGGGGGGGGGGGGG
CTGGGGCTACATGTGAAGGACAGGACATTTCAGCTGGCTGCACTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
TAGTTCTCCATTCACCCCTGAGCCAGCTGGGGCCACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
CCTCTGCTGCTGCCCTGCTGTGAGGTAAGGTGGGGAGGGGGCTACAGCCCACATCCCACCCCTGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTCACTGGGGCTCAGTCCAGGCTGACTTCCAACTACAGCATTGACCCCAACTTGG
TGTGAAGGAAAGGAAAGGATCTGGGCTTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG
CTTGGGGAGTGCAGGGGGCTGGGGCTGGGGCTGGGGCTGCACAGAACGGGGGGGGGGGGGGGGGGGGGGGGGGGG
GGCCTGGACCCCTGTCAGGATGGGGCCACCTCAGAACCAAACACTCAGTCCCCACTGTGGCATGAGGGGGGGGG
CCATGTTGAGGGCGAAGGGCAAGGGCTTGTGTGTCTGGGGAGGGAAAGGAAAGGGTGTGGAGGGCTTAATTATGG
ACTGGGGAAAAGGG
GAGCAGGAGCTGGCTGACTGCTCAGACTCTGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG
GGGACCAAGCCCAACTTCCACTTGTGACTGCTGGGGACCTATTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
AACTATTGTTATAGTTGTTAATTAATAGCTGTGATTTCAAGGTTCTTGTGATTTGAAACAGAGTTATGCT
GATTGTAACCTCCCAAGCCGCCAGTGGGATGGGGAGGGAGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
GTCCAGAGAAATTCCCTTGGGACTGGAGGAGGAGAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
GTTGGGGAGGGGGCTGG
AACCTGAAGGTTGGAACTCAGGTTATTCCTGCGCTGG
TGAGGGTGGCTCTACTGTGAATTGG
CTTCTGCTGCCCTGCTTAACAATGCCGGCCAACCTGGGACCTCACGGTTGCACTCCATTCCACCAAGATGACCTGA
TGAGGAAATCTCAATAGGATGAAAAGATCAATGAAAATTGTTATATGAAACATATACTGGAGTCGTAAAAAG
CAAATTAAAGAAGAATTGGACGTTAGAAGTTGTCAATTAAAGCAGCCCTAATAAAAGTTGTTCAAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHVLVAGGG
AGAVSRTCTAPLDRLKVLQMVKHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDILAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATA CGCTGGAACCTTCCCCAGGCCTGGC
TTCCCTGGGGCAGATCCTCTTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGTTAGGAGACACTCCATCACAGTCACACTGTGCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTGAACCTGACATCAAACCTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTAGGCTGAGCCAGCAGTGTTGCTGATCAAAGTGTAGTT
GGCAATGCCCTTTGGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAAACTGGAGCCTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTCCCC
CAGCCCACAGTGGCTGGGCATCCAAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACTCTGAGAATGTTGACCATGAAGGTTGTCTGTGCTCTACAATGTTACGA
TCAACAAACACATACTCCTGTATGATTGAAAATGACATTGCCAACAGGGATATCAAAGTG
ACAGAACATCGGAGATCAAAAGGGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTGT
CTCTCTTCTTGCCATCAGCTGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAATAAT
GTGCCTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCAC
CACCAAGATATGACCTAGTTATTTCTGGAGGAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAACAAAAAGCAGCAAGGCTCAATATGAACAAGATAAAAT
CTATCTCAAAGACATATTAGAAGTGGAAAATAATTGATGTAAGTGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGGATAGTGCATGTTCTGTGCTCTGAATTAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCCTGAAAGTCTATCCAAACATATCCACATCTTATATCCAC
AAATTAAGCTGTAGTATGTACCCCTAACAGCCTGCTAAAGTGCATGCTTCCAAAGGTGCCTGGCTTC
TCTTCCAACTGACAAATGCCAAAGTGGAGAAAATGATCATAATTAGTACATAACAGAGCAGT
CGGGGACACCGATTTATAAATAAACTGAGCACCTTCTTTAAACAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCKMIENDIAKATGDIKVTSEIKRSHLQLLNASKL
CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCTAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCTGGTGGTTATGGGGTCTGCCCTGGAGGGCTTATGGACCACAGCTGG
TGGAGGGCCATGGACACCCAATCCTGGATGTTCCCTCTGGAACCTCAGGAGGACCATATG
GCGGTGCAGCTCCGGGGCCCTATGGTCAGCCACCTCAAGTCTTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGCCCTCCAAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCAGGCTCTCAGCCCTGTGAAATTCATCCAGCAGTGGAAAGAACCTT
CCAGCAGTATGACCGGGACCGCTCGGCTCATTAGCTACACAGAGCTGCAGCAAGCTCTGCTCC
AAATGGCTACAACCTGAGCCCCAGTTCACCCAGCTTCTGGTCTCCGCTACTGCCAACGCTCT
GCCAATCCTGCCATCGAGCTTGACCGCTCATCCAGGTGTGCACCCAGCTGCAGGTGTCAGAGA
GGCCTTCCGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTCGAGGACTTCGTC
CCATGACAGCTCTCGGATGCTGACCCACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTAGAGTGAGAGAAGTATGTGGACATCTTCTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTGCTGTGAAACACTGTTCAAAGAGGGTGGAGAGTCCTGCATCATAGCCACCA
AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCGTATGCCAGTGGTAGTGTCTCATCGGCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAAGGCCATCTGTCAAACGAGGCCATTCTCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAAGTGGCTGGATTCTGCCACACCCATAAAAT
CCTTGTGTGTTAACTCTAGCTGCCTGGGCTGCCCTGCTCAGACAAATCTGCTCCCTGGGCA
CTTGGCCAGGCTCTGCCCTGAGCTGGACCCCTACTGCCATGCTGCTGCC
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTAATTGCAATTTC
ATTGGGGCAAAAGTCAGTGAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSTPGGPYGGAAPGGPYGQPPPSSYGAQQPGLYGQGGAPPNDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTSGRIDVYGFSAWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTTTGCCCGAGCATCGGTGTTGGAGGAGAAAGTTCCAAAACCTCGGACC
AACTTGCTCAGCTCGGACAACCTCCTCACTGGCCCTCTAACACTGAACATCCGCAGCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGAGGTGGCCTCCATCGTGGGGCTGCCTGCCATG
GATTCCCTGGCCCCCTGAGGATCCTGGCAGATGATGGCTGCTGCGCTGAGGACCGCCTGGGGA
AGCGCTGCCTGAAGAACTCTTACCTCTCCAGTGTGCAGGCCCTCGCTCGGCAGTGGCCCTT
TGCCTGGGAGTCTCTCCGATGCCACAGGCCTCTCACCTGAGGTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAACACTGGGAGCCGGGGAAAAATCCTTCCAAACG
CCCTCCCTGGTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGTACCTGAATCCCAGTG
TGTCCCTGGGAGGGTGGAGGCCCTGGGACTGGTTGGGAAACGAGGCCATGCCACACCCCTGAGGGA
ATCTGGGTATCAATAATCAACCCCCAGGTACCAAGCTGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGAAATATTAATCGGTATCCAGGAG
GCAGCTGGGAAATATTCACTATACCCAGGTATCAATAACCCATTCCCTGGAGTTCTCCGC
CCTCCTGGCTTCTTGGAACATCCCAGCTGGCTCCCTAACCTCCAAGCCCTAGGTGAGTG
GGGCTAGAGCACGATAGAGGGAAACCCAACATTGGGAGTTAGAGTCCTGCTCCGCCCTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCCACTTTCACTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMHPPEGIWGGINNQPPGTWGNINRYPGGS
WGNINRYPGGSGWGNINRYPGGSGNIHLYPGINNPFPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCATCCCAGGAGGCAGTGGCCACTATGGGTC
TGGGCTGCCCTTGTCCCTCTTGACCCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTCATGAGTCCAGCTCCTGGAA
TTGCTTAAAAGCTCTGCCTCCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTAACACCATGTTGTCGAACACATTGACAGCCATTGAAGCCTGTGCCTTCTTGGCC
GGGCTTTGGCCGGGATGCAGGAGGCAGGCCCCGACCCCTGTCTTCAGCAGGCCCCCACCTC
CTGAGTGGCAATAAATAAATTCTGGTATGCTG

FIGURE 66

MGSGLPLVLLLTLGSSHGTGPGMTLQLKLKESFLTNSSYESSFELLEKLCLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGGC
CAGGTGCCCGTCCGAGGTGCCCTGGCCGGAGATGCCGTAGGAGGGCGAGCGCGAGAAGCCCC
TTCCTCGCGCTGCCAACCGCCACCCAGCCCATGGCGAACCCGGCTGGGCTGCTTCTGGCG
CTGGGCTGCCGTTCTGCTGCCCGCTGGGCCGAGCCTGGGGCAAATACAGACCACCTCTGC
AAATGAGAATAGCACTGTTGCCTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGCTTCTCCCTCTGGCTGCCTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGCGGAAGCTTCGGAGAACGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCATCAGGTCCCCTCCTGCATCTGTCTCCCTATTGCTGTGACCTTGGGAAA
GGCAGTGCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATATTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAA

FIGURE 68

MANPGLLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGTTATTCAGAGGTTTGT~~TTTCTT~~TAGTCTGTGCCTGTCACCAG
TCAAATAC~~TT~~CTCATTAGCTGAATAATAATGGCTTGAAGATATTGT~~CATT~~GTATAGATCCTAGTGTGCCAGAA
GATGAAAAAATAATTGAACAAATAGAGGATATGGT~~GACTACAGCTCTACG~~TCTACGCTACCTGTTGAAGCCACAGAAAAAGA
TTTTTTCAAAATGTATCTATATTAA~~TT~~CCTGAGATTGGAGGAAATCCTCAGTACAAAAGGCCAAAACATGAA
AACCATAAACATGCTGATGTTATAGTGCACCACCTACACTCCAGTAGAGATGAACCATAACCAAGCAGTTACA
GAATGTGGAGAGAAGCGAATA~~CAC~~TGGGAGGCTACACTGGGAGGAGTGTGATGACTACATGGACCA
CCAGGCAACTGTTGTCATGAGTGGGCTACCTCCGGTGGGAGTGTGATGACTACATGGAGATCAGCCTTC
TACCGTCTAAGTCAAAAGCAAGGCAAGGTTCCGAGGTATCTGGTAGAAATAGAGTTATAAGTGT
CAAGGAGGCAGCTGCTTAGTAGAGCATGCAGAATTGATTCTACACAAAAGTGTATGGAAAAGATTGTCAATTCTT
CCTGATAAGTCAAAACAGAAAAGCATCCATAATGCAAAGTATTGATTCTGTTGTGAATTGTAAAGAA
AAAACCCATAATCAAGAAGCTCAACCCCTACAAAACATAAAGTCAATTAGAAGTACATGGGAGGTATTAGCAAT
TCTGAGGATTAAAACACCATACCCATGGT~~GACACACCT~~CTCACCCTGTCATGCTGAAGATCAGTC
AGAATTGTCGTTAGTTCTGATAAGTGTGAAAGCATGGGGT~~PAGGACGCC~~AAATCGAATGAATCAAGCAGCA
AAACATTCTGCTGCAAGTGTGAAAATGGATCTGGGGATGGTCACTTGTAGTACTGCCACTATTGTA
AATAAGCTAATCCAATAAAAGCAGTGTGAAAGAACACACTCATGGCAGGATTACCTACATATCCTCTGGAGGA
ACTTCCATCTGCTGTAATTAGTCATTCAGGTGATTGGAGAGCTACATTCCAAC~~TG~~ATGGATCCGAAGTA
CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTCTTGATGAA~~AGT~~GAACAAAGTGGGCCATTGTCAT
TTTATTGCTTGGGAGAGCTGCTGATGAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCA~~TTT~~TATGTT
TCAGATGAGCTCAGAACAA~~TGGC~~CTCATTGATGCTTTGGGCTTACATCAGGAAATACTGATCTCCAGAAG
TCCCTTCAAGCTGAAAGTAAGGGATAACACTGAATAGTAATGCCGGATGAACGACACTGTCATAATTGATAGTACA
GTGGGAAGGACAGCTTCTCATCACATGGAACAGTCTGCCCTCCAGTATTCTCTG~~GG~~ATCCCAGTGGACAA
ATAATGGAAAATTCACAGTGGATGCAACTGGCTTCAAGGAACTGCAAAAGTGGCACT
TGGGCAATACAATCTCAAGC~~CC~~AAAGC~~CC~~AAAGC~~CC~~AAAGC~~CC~~AAACTTCCAGTAACTGTCAGCAGCAAATTCTCT
GTGCCCTCAATCACAGTGAATGCTAAATGAATAAGGAGCTAAACAGTTCCAGCCAATGATTGTTACGAGCAA
ATTCTACAGGATATG~~T~~AC~~T~~GTTCTGGAGCCAATGTGACTGCTTCATTGAATCAGAA~~T~~GGACATACAGAAGTT
TTGGAACTTTGGATAATG~~T~~GGCAGCGCTGATTCTTCAAGAATGATGGAGTCTACTCCAGTATTACAGCATA
ACAGAAAATGGCAGATAAGCTTAAAGTTCAGGCTCATGGAGGAGCAACACTGCCAGGCTAAATTACGCCCTCA
CTGAATAGAGCCCGTACATACAGCCTGGTAGTGAACGGGGAAATTGAAGCAACCCGCCAACCTGAAATTGAT
GAGGATACTCAGACCCACCTGGAGGATT~~T~~AGCCGAACAGCATCCGGAGGTGATTGTTG~~T~~ATCAGTCCAGC
CTTCCCTGCC~~T~~GA~~CC~~AATACCCACCAAGTCAAATCACAGACCTGATGCCACAGTTCATGAGGATAAGATTATTCTT
ACATGGACAGCAC~~CC~~AGGAGATAATTGATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGAAGTATTCTT
GATCTAAGAGACAGTTGATGATGCTCTCAAGTAATACTACTGATCTGTCAACAAAGGAGGCCAAC~~T~~CCAGGAA
AGCTT~~T~~GCATTAAACCAAATATCTCAGAAGAAAATGCAACCCACATATTG~~C~~ATTAAAGTATAGATAAAA
AGCAATTGACATCAAAGTATCCACACATTGCAACAGTAAC~~T~~TTCTTATCCCTCAAGCAAATCTGATGACATTGAT
CCTACACCTACTCTACTCCACTCTACTCTGATAAAAGT~~C~~ATAATTG~~G~~AGTTAATTCTACGCTGGTATTG
TCTG~~T~~GATTGGGCTGTTGTAATTG~~T~~TAACTTTAAGTACCCACATT~~T~~GAACCTTAACGAAGAAAAAAATCTC
AAGTAGACCTAGAGAGGTTAAAACAAAACATGTAAGTAAGGATAATTCTGAATCTTAAATTCATCCCAT
GTGTGATCATAAAACTCATAAAAAATTAAAGATGTCGGAAAAGGATACTTTGATAAAATAAAACACTCATGGATA
TGTAAGGAAACTGTCAAGATTAAAATTAAAGTGTATTCTGTTATTG~~T~~ATTTG~~T~~ATAAGGAAATAGTGTGAAACAG
ATCCTTTTCATACTGATAC~~T~~GGTGTATTATTG~~T~~ATGCAACAGTTCTGAAATGATATTCAAA~~T~~GCATCAA
GAAATTAAAATCATCTATGAGTA~~C~~AAACATAACTAAAGGAGGAAATAACACATTGGAAAAA
AAA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTYLF
ATEKRFFFKNVSILI PENWKENPQYKRPKHENKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLGGKKQNEYGPPGKLFVHEWAHLRGVFDEYNEDQPFYRAKSKKIEATRC SAGISGRN
RVYKCQGGSCLSACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDS VVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVT PPPPVFSLLKISQRIVCLVLDKGSGMGKDRLNR
MNQAAKHFLQLT VENG SWGMVHF DSTATIVNKLIQIKSSDERNTL MAGLPTYPLGGTSICSGIK
YAFQVIGELHSQDGSEVLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDA FGALTSGNTDLSQKSLQLESKGTLNSNAWMNDTVIIDSTVGKDFFL
ITWNSLPPISLWDPSGTIMENFTV DATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKN DGVYSRYFTAYTENGRYSLKVRAGGANTARLKL RPPLNRAAYIPGWVVNGEIEANPP
RPEIDEDTQTTLED FSRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN
FDVGVKVQRYIIRISASILDRLDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPKSHNSGVNISTLVLSVIGSVVI
VNFILESTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAAGACCATACGTCCCCGGGCAGGGTGA
CAACAGGTGTATCTTTGATCTCGTGTGCTGCCTTCATTCAAGGAAAGACGCCAAGGTAAAGGACCCA
GAGGACCAATGATGTAGCCACCTCTAACCTTCCCTCTGAACCCCCAGTTAGCAGGATTTACTAGAGGTGTC
ACTCAACCAGCAAGCGCTCTCGGCTTAACTTGTTGGAGAGAACCCTTGTTGGGGCTCGCTCTTAGCA
GTGCTCAGAAGTGACTTGCCTGAGGGTGGAGAGAAGAAAGGTTCCCTCTGCTGTTGGCTGCACATCAGGAA
GGCTGTGATGGAAAGGTGAAAATGGAGATTCACCTTCAGTCTGCTTCAGGAGATCCTTAAAGTCAAGGAA
AGTAGAGAAGCTCTCTGTTGTTGTTAACCTAACAGGAGCAGAACCTGTTAGAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCAAACGCATGCTCTGTTGCTAGCCAGGGAAGCCCTTCGTTGGGGCCCGCTTGGGGATGCC
ACCGGTTCTGGACGCTGGCTGATTCTGTAATGATGATGGTTCGCCGGGGCTGCTGCTGTTGGGGATTTCCCGGGTGGT
GTTTCTGTTGCTCTCTGCTGCTGCTATCTGTTCTGTAATGTTGGCTGACCCCAAAGGTGACGAGGACAG
CTGGCACTGCCAGGGCAACAGGGCACCCACGGGAGGGGGTACAGGGCTCTCAGGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCTGAAGGGCAGATGCCAACGCTCAAGGAGGACTGCAAGGAGGACTGAGCAGCTCAGG
AATGGCAGTACCAAGCCAGCAGTCGCTGCCCTGGCTCTGGACAGGAGCCCCAGAGAAAACCCAGGCCACCTC
CTGGCCTCTGCACTCGCAGGTGGACAAGGAGGTAATGCTGGCTCAAGCTGGCACAGAGTATGAGCAGTG
CCTTTGAGTCTTAACTCTACAGAAGGTGACAGCTGGAGACTGGCTTACCGCCACCCGAGGAGAACGCTGTG
AGGAAGGACAAGGGATGAGTTGGAGACATTGAATCAGCCTTGGAGACCTGAACAAATCTGAGAGAACAGC
CCCAATCACCCTCTAACAGGCTCTGATTCTGAGAAGGAGATCAGCAGAACAGAAGGGACAATGGACATTGTAT
GAGCTCACCTAACAGGAGCACAAACAGAACCTAACAGGCTCATCTTATTGACCATTCAGCCCCATCATGAAA
GTGAAAAATGAAAGCTAACATGGCAACACGCTTACATGTTATGTCAGTCTAGCAAAAGGGTGGACAAGTTC
CGGCAGTCATGAGAATTTCAGGGAGATGTCATTGAGCAGGATGGGAGACTCCTACTGTTGTTACTTGGG
AAAGAAGAAATAATGAACTTCAAGGAAACTTGAAGGACTTCAAAAGCTGCAACTTCAAGGACTTACCTTCATC
CAGCTGAATGGAAATTCTCGGGAAAGGGACTGATGTTGGAGCCCTCTGGAAGGGAGGAACGTCCTCTC
TTTTCTGATGTCGACATCTACATTCACATCTGAATCTCTAACATACGTCAGGCTGAATACACAGCCAGGGAAAG
GTATTTATCCAGTTCTTCAGTCAGTACAATCTGGCATATAACAGGCCACCATGATGCACTCCCTCCCTGGAA
CAGCAGTCGTTCATAAAGAGGAACTGGATTGGAGAGACTTGGATTGGGATGACGTGTCAGTATGGTCAGAC
TTCATCACATATGGTTGTTGATGGACATCAAAAGGCTGGGGAGGAGTGCACCTTATCCAAGTATCTC
CACAGCAACCTCATGTCAGGAGGACTCTTCACCTCTGGCATGAGAACGGCTCATGGAGCAG
CTGACCCCCGACAGTACAAGATGTCAGTCAGGCTCAAGGCCATGAGAACGGCATCCACGGCAGCTGGCATCTG
GTGTTAGGACAGAGATAGGGCTCACCTTCGCAACAGAAACAGAACAGAACAGTAGAACAAAAACATGAACTCCAGA
GAAGGATGTTGGAGACACTTTTCTTCTTGCATTAACAGTGAAGTGGCTGCAACAGAGAACAGACTTCCATAAA
GGACGCAAAAGAAATTGGACTGATGGTCAGAGATGAGAACGGCTGGGATTTCTCTGTTGGCTTTTACAACAGA
AATCAAAATCTCCCTGGCTGCAAAAGTACCCAGTGGACCTCTGTAAGTGTGACAAAGGCAGATGCTGTG
AGATTATAAGCTTAATGGTGTGAGTTGATGGTTTACATACGTCAGGCTGTTGCTGCTATTGATG
AATATTGATTTAAAGAGCAGTTTGTAAAAATTCATAGCATGAAAGGCAAGCATATTCTCTCATATGAATGA
GCCTATCAGCAGGGCTAGTTCTAGGAATGCTAAATATCAGAAGGCAAGGAGGAGATGGCTTATTGATGACT
AGTGAGTACATTAAGTAAATAAAATGGACAGAACAGAACACCATATACTGTCATATTCCCCAAAGAT
TAACAAAAATAATGTTATCTTGGTCTTTAACCTGTTACTGTCCTTTTCTTAAAGGAAACTGACT
TTTTCTCTGTTGAGTTAGTCTGTTATTAATACCTTGCACGCCCTAACAGGACAAAGGACACAAGTGGCTAC
ATTTTATATTTTAAGAAGATACTTGAGATGCAATTGAGACTTCAAGGCTAACAGCATCAAATGATGCCATAT
CCAAGGACATGCCAATGTCATTCTGTCAGGCACTGAATGTCAGGCATTGAGACATAGGGAAGGAATGGTTGACT
AATACAGACGTCAGAGATACTTCTGTAAGAGATTTTCGAAGAGGAGCAACTGAAACACTGGAGGAAAGAACATGAC
ACTTTCTGCTTACAGAAAGGAACTCATTGAGACTGTCAGTGTGATATCTGAGTCACTAAAGTCAGAACACATTT
CTCCTCAGAAGTAGGGACGCCCTTCTACCTGTTAACAAAGGATACCTGTTGAGACAAACATCTCTTCA
AAAACAGGGTGTCTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATTGT
GAAAGATCAATCCATCTGCGAGAATCTAGTGGGATGGAAGTTTGCTACATGTTATCCACCCAGGCCAGGTGGAG
TAAGTCAATTGTTAAATTAAGCAGTTCTACTCAATCAGAACGATGCTTCTGAAATTCATTTTATTACCATTT
CAAACATTTTAAAGGAAACTACAGTAAACATAGTGGTTCTCTGTTGCTTGTACAGTAAACTGTTTAAAGCTTCAAGAAC
ATGATGAGCTAATTATCTTTGAGTCCTGCTTCTGTTGCTCACAGTAAACTGTTTAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTGTTAAAAATGCACTGTTGATTGATTGACTGGTAGTTGATGAAATTAAACACAGG
CCATGAATGGAAGGTGGATTGCAAGCTAATAAAATGATTGTTGGATATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRPEEKPVRKDRDELVEAIESALETLNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVNEKLNMAN
TLINIVPPLAKRVDKFRQFMQNFRMCEIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQILNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFQSQY
NPGIIYGHHDADVPPLEQQQLVIKKETGFWRDFGFGMTQYRSDFINIGGFDDIKGWGGEDVHLYR
KYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPEQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTGTCCTGGGATCCA
GAAACCATGATACCCTACTGAACACCGAATCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCACTCC
CCTCCCTCTCTCTGCCTGTCCTAGTCCTCTAGTCCTCAAATTCCCAGTCCCCTGCACCCCTC
CTGGGACACTATGTTGTTCTCGCCCTCCTGCTGGAGGTGATTGGATCCTGGCTGCAGATGGG
GTCAACACTGGACGTATGAGGCCACATGGTCAGGACCATTGCCAGCCTTACCCCTGAGTGT
GGAAAACAATGCCAGTCGCCATCGATATTCAAGACAGACAGTGTGACATTGACCCCTGATTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCAGCCTTGGACCTGCACAACAATGCC
ACACAGTCAACTCTCTGCCCTCTACCCCTGTATCTGGTGACTTCCCAGAAAATATGTAGCT
GCCAGCTCCACCTGCACTGGGTCAAGAAAGGATCCCCAGGGGGGTCAAGAACACCAGATCAACAG
TGAAGCCACATTGCAAGAGCTCCACATTGTACATTATGACTCTGATTCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGCCCTAGGCCCTGGCTGTCCCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGTTATGAACACATTGAGTCACCTGCATGAAGTCAGGCATAAAAGATCAGAACACCTC
AGTCCTCCCTCAACCTAACAGAGAGCTGCTCCCCAACAGCTGGGCAGTACTCCGCTACAATG
GCTCGCTCACAACTCCCCCTGCTACCGAGCTGTGCTCTGGACAGTTTTATAGAAGGTCCCAG
ATTTCAATGGAACAGCTGAAAAGCTTCAGGGACATTGTTCTCACAGAAGAGGAGCCCTCTAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGCATGGTCTTGCTTCTTCA
TCCAAGCAGGATCCTGTATACCAACAGGTGAAATGCTGAGTCTAGGTGAGGATCTGGTGGC
TGTCTCTGCCTCTCCTGGCTGTTATTCATTGCTAGAAAGATTGGAAGAAGAGGCTGGAAAA
CCGAAAGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCATAAAATCCTCTCAGATAC
CATGGATGTGGATGACTCCCTCATGCCATCAGGAAGCCTCTAAATGGGTGAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCCTGGACATCTCTAGAGAGGAAT
GGACCCAGGCTGTCACTCCAGGAAGAACTGCAGAGCCTCAGCCTCTCAAACATGTAGGAGGA
ATGAGGAAATCGCTGTGTTAATGCAGAGANCAAACCTGTTAGTGCAGGGGAAGTTGGG
ATATACCCCAAAGCCTCTACCCCTCACTTTATGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTAGGATAAAGAGTTGCTGTTGAAGTTGTATTTTGATCAATATATTGAAATTAAG
TTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSVTFPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCCCTGGGACGGCAGTCCCTGTGTC
TCTGGGGTTGCCTAAACCTGCAAACATCACCTTCTATCCATCAACATGAAGAATGCCTACA
ATGGACTCCACCAGAGGGCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTCATCACAA
ATTGGCCCACCAGAGGGTGCACTGACTACAGATGAGAAGTCCATTCTGTTGTCCGTGACAGCTCC
AGAGAAAGTGGAGAGAAAATCCAGAACAGCCTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTGTGAATACTAAACAGAACAGTGGTCCAGTGTGTGACCAACCAC
ACGCTGGTGCTCACCTGGCTGGAGGCCAACACTCTTACTGCGTACACGTGGAGTCCGTCCC
AGGGCCCCCTCGCCGTGCTCAGCCTCTGAGAACAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTGCCCCATATCTATTACCGTGTGTTCTTTT
TCTGTGATGGGCTATTCCATCTACCGATAATCCACGGTGGCAAAGAGAAAACACCCAGCAAATT
GATTTGATTTATGGAAATGAATTGACAAAAGATTCTTGTGCCTGCTGAAAAAAATCGTGTGATTA
ACTTTATCACCTCAATATCTGGATGATTCTAAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCACTGATGATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGTATGCTCGCATTGATGGAAATTGGTGACTCTGAAG
AAAACACGGAAGGTACTTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCAACTGACATTGCGGGGCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGAGGTGCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTGGCAG
TCTTGGGCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCTGGCGACTGGGA
TCCCCAAACTGGCAGGCTGTGTATTCCCTCGCTGTCCAGCTTCGACCAGGATTAGAGGGCTGCG
AGCCTCTGAGGGGATGGCTCGGAGAGGGAGGGCTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAAGGAGAAAATGAAACCTATCTCATGCAATTCTATGGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGGCAACACTCCCTTGCCTTGTGCAAACAAGTGAG
TCACCCCTTGATCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTCCAGTTGTCACTG
CTGTGAGAATTACTTATTCTCTATTCTCATAGCACGTGTGATTGGTTCATGCATGTA
GGTCTCTAACATGATGGTGGCCTCTGGAGTCCAGGGCTGGCCGGTTGTCTATGCAGAGAA
AGCAGTCATAAATGTTGCCAGACTGGGTGCAGAATTATTACAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTPKSNRWSQLCVTNHTLVLTWLEPNLTYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDDKRFFVPAEK
IVINFITLNISDDSKISHQDMSSLGKSSDVSSLNDPQPSGNLRPPQEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLOQEEVSTQGTLLESQA
ALAVLGPQLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLLSSFDQDS
EGCEPSEGDSLGEEGLLSRLYEEPAPDRPGENETYLMQFMEEWGLYVQMN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGGGGGCCGAGGATCCAGCGTGCCAGGTCTGGCATCCTGCACACTTGCTGCCCTCTGACAC
CTGGGAAGAGATGGCCGGCCCTGGACCTTCACCCTCTCTGTGGTTGCTGGCAGCCACCTTGATC
CAAGCCACCCTCAGTCCCACTGCAGTCTCATCCTCGGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCTGTGCTGGCAGCTGGTGAACACCGTCTGAAGCACATC
ATCTGGCTGAAGGTCATCACAGTAACATCCTCCAGTGCAGGTAAGCCCTCGGCAATGACCA
GGAGCTGTAGTCAAGATCCCCTGGACTGGTGGCTGGATTCAACACGCCCCTGGTCAAGACCA
TCGTGGAGTCCACATGACGACTGAGGCCAAGCCCCCATCCGCATGGACCCAGTGCAGTGGC
CCCACCCGCTGGTCTCAGTGACTGTGCCACCAGCCATGGGAGCTGGCACTCACTGTGA
TAAGCTCCTCCTGGTGAACCCCTTAGCTAGCAGGTCATGACCTCCTTAGGCCACTCCTGC
CCAACTTAGTGAAAACCAGTGTGCCGTGATGAGGCTCTTCTAAGGCATTATGCA
CTCCTGCAGTGGTGAAGGTGCCATTTCCTCAGCATTGACCGTCGGAGTTTGACCTTCTGTATCCTGCATCAGGGTGACACCATTCAGCTCACTGGGGCCAAGTTGGGACTCAGGGAA
AGGTGACAGTGTTCAAAACTTGCGAGTCCCTGACAATGCCACCCTTGACAACACTCCCG
TTCAGCCTCATCGTGAGTCAGGAGTGGTAAGGCTGCAGTGCTGCTGCTCCAGAA
ATTCATGGTCCTGTTGACTGTGCTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGG
TGATCAATGAAAAGGCTGCAGATAAGCTGGATCTACCAGTCGTGAAAGTCTAACTCAGG
ACTCCCGAGTTTTTTTATAGACCAGGCATGCCAGGTGGCCCAACTGATCGTGCTGGAAGTGTT
TCCCTCCAGTGAAGCCCTCCCCCTTGTTCACCCTGGCATGAAGCCAGTGCTGGAAGTCAGT
TTTACACCAAAGGTGACACTTATACTCACTGAATAACATCAGCTCGTGATGGATCAGGTG
ATGAACTTGGGATTGGCTGGTTCAACCTGATTGTTCTGAAAAACATCATCACTCAGGATCATCCAA
CTCCCACTCGTGCCGAACCAAGTGGAAATTAAGACTTGGGTCCAGTGTCATGGTGAAGG
CCTGGGATTCGGAGGCAGTGAGTCCTCATGACCAGGATGCCTGTGCTTACCAGCCTC
TTTGTGAAACCCAGCTCCTGTCCCAGTTGAAGTTGGGATGGCAGCCATCAGGGAGGCTGG
GTCCCCAGTTGGGAGTATGGTGTGAGCTCATAGACCATCCCCTCTCTGAATCAAAACATTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPMLDNIPFSLIVSQDVVKAAVAAVLSPPEFMVLLDSVLPEAHRLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSEALRPLFTLGIASSEAQFYT
KGDQLIILNLISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTPASLWKSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCTGAAGACGCTTACTGAGAGGTCTGCCATGGCCT
CTTGGCCTCCAACTTGTGGCTACATCCTAGGCCTCTGGGCTTTGGGCACACTGGTGC
GCTGCTCCCAGCTGGAAAACAAGTTCTTATGTCGGTGC
CACATTGTGACAGCAGTTGGC
CCAAGGGCTCTGGATGGAATGTGCCACACACAGCACAGGCATACCCAGTGTGACATCTATAGC
ACCCTCTGGCCTGCCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTTGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGTAGCAGGTGGAGTCTTTTCA
TCATCCTTGAGGCTCTGGGACTTCTACTCACCAGTGGTGC
CATGAAATTGAGATTGGAGAGGCTCTTACTTGGCATTATTC
CTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTCC
TGCTCATCCCAGAGAAATCGCTCCA
ACTACTACGATGCC
TAC
CAAGCCAAACCTCTGCCACAAGGAGCTCCA
AGGCCTGGTCAACCTCC
AAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTTGAAGAACCAGGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCC
ACAGGTGAGGGACACTACCAACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTGGCATTGGATTGAGCAAAGGCAGAAATGGG
GCTAGTGTAA
CAGCATGCAGGTGAATTGCCAAGGATGCTGCC
ATGCCAGCCTTCTGTTCC
TCACCTTGCTGCTCCCTGCC
CTAAGTCCCCAACCTCA
ACTTGAAACCCATTCC
TTAAGCCA
GGACTCAGAGGATCCCTTGCC
CTGGTTAC
CTGGACTCC
ATCCCCAACCC
ACTAATCACA
TCCC
ACTGACTGCC
CTGTGATCAAAGACCC
CTCTCTGGCTGAGGTTGGCT
CTTAGCTCATT
GCTGGGGATGGGAGAGAAGCAGTGG
CTTGTGGCATTGCTCTAAC
CTACTTCTCAAGCTTC
CCTCC
AAAGAAA
CTGATTGGCCTGG
AACCTCC
ACTCTTGT
TATGACTCC
ACAGTGTCCA
GACTA
ATTGTG
CATGAA
CTGAA
AAACCA
CTC
ACGGT
ATCC
AGGG
AACAG
AAGCAG
GGAT
GGGAGG
ACAGGA
AGGCAG
GCC
CTGG
GACATT
AAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAQAMMVTTSSAISSLACIIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLLHGIILRDFYSPLVPSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCCACCGCGTCCGCGCCTCTCCCTTGTGGACCTTCCTCGTCCTCCATCTCTCCCTCCTTC
CCCCCGTCTCTTTCCACCTTCTCTTCCCACCTAGACCTCCCTGCCCTCCCTTCT
GCCCAACCGCTGCTCCCTGGCCCTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGCTGTGG
GTTGATCTGTGGCCCCGTGCCTCCGTGTCTTCTCGTCCTCCCTCCGACTCCGCTCCCG
ACCAGCGGCCTGACCCCTGGGAAAGGATGGTCCCGAGGGTAGGGTCTCTCCCTCTTGCTGGGA
CTCGCGCTGCTCTGGTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACATGTTCTGCCTTT
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGGCCACAAGGCTGATGT
ACTGCCTGCGCTGTACCTGCTCAGAGGGGCCCATGTGAGTTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCTCAGCCTGTGACGGAGCCACAGCAATGCTGTCCTGAGCTGTGTGGAACCTCACAC
TCCCTTGGACTCCGGGCCCCACCAAAGTCTGCGCAGCACACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCGCTGCCAACCAAGTGTGTCTGAGCTG
ACAGAGGGCCAGATCTACTGCGGCCTCACAAACCTGCCCGAACCAAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCTGCTGCCAAGCCTGCAAAGATGAGGAAGTGAAGCAATGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCAAGTGTGTGCTGGAGAAAAG
AGAGGCCGGCACCCAGCCCCACTGCCCTCAGGCCCTCTGAGCTTCATCCCTGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAGATGTCCTGAAGGAGAAACATAAGAAAGCT
GTGTGCATGGGGAGACGTACTCCCACGGGAGGTGTGGCACCCGGCTTCCGTGCCCTCGGC
CCCTTGCCCTGCATCTATGCACCTGTGAGGATGCCGCCAGGACTGCCAGCGTGTGACCTGTC
CACCGAGTACCCCTGCCGTACCCGAGAAAGTGGCTGGAAAGTGTGCAAGATTGCCAGAGG
ACAAAGCAGACCCCTGCCACAGTGAGATCAGTTCTACCAAGGTGTCCCAAGGCACCGGGCCGGGCT
CTCGTCCACACATCGGTATCCCCAAGCCAGACAACCTGCGTCGCTTGCCTGGAACACGAGG
CTCGGACTTGGTGGAGATCTACCTCTGGAAAGCTGGTAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCAAGGCCACAGCCAGAATCTTCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTCCAGAAAGAGGCACAGCACTCCGACTGCTCGTGGCCCCACGAAGGTCACT
GGAACGTCTCCCTAGGCCAGACCCCTGGAGCTGAAGGTCAAGGCCAGTCCAGACAAAGTGAACAG
ACATAACAAAGACCTTAACAGTTGCAGATATGAGCTGTATAATTGTTATTATATATTAA
TAAGAAGTTGCATTACCCCTCAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELP
PSRLPNQCVLCSCTEGQIYCGLTTCEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGT PAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKKACVHGGKTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLSDQESQEARNPERGTALPTARWPPRRSLERLPSPDPAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCGCTCACGCAGAGCCTCTCC
GTGGCTCCGCACCTGAGCATTAGGCCAGTTCTCCTCTCTAATCCATCCGTACACCTCTCCTGTCA
TCGGTTCCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTGGTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGGCCAGACAAGCCTGTCCAGGCCTGGTGGGGAG
GACGCAGCATTCTCCTGTTCCGTCTAAGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAACTGGTGAAGGATTCTATTGCGGAGGGCGCATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCCTATGGTGAGGATTAGTTCCAGTCTTATACACAGAACGCCATCTGGAGCT
ACAGGGTGTCAGCACTGGGCTCAGTCTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGCCGTGTTGATGTGGAGATCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCCGATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATAACCTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGAAACTCTGCTGTGGCTA
TTTTTGGCATTGTTGACTGAAGATTTCTCTCAAATTCCAGTGGAAAATCCAGGCCAACTGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGAAACACGCAGCTGGAGGTGACTCTGGATCCAG
AGACGGCTACCCGAAGCTCGGTTCTGATCTGAAACTGTAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCTGGGAGTGTGCCGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTGTCTCCGATCATGGTACTGGTCTCAGACTGAATGGAGAACATTGTATTC
ACATTAAAATCCCGTTTATCAGCGTCTCCCCAGGACCCACCTACAAAATAGGGTCTTCTGGACTA
TGAGTGTGGGACCATCTCTTCTCAACATAAAATGACCAAGTCCCTTATTTATACCCTGACATGTCGGTTTG
AAGGCTTATTGAGGCCTACATTGAGTATCCGCTTATAATGAGCAAATGGAACTCCCATAGTCATCTGC
CCAGTCACCCAGAATCAGAGAAAGGCCCTCTGGCAAAGGCCCTCTGCAATCCCAGAGACAAGCACAG
TGAGTCCTCTCACAGGCAACCACGCCCTTCTCCCCAGGGTGAAATGTAGGATGAATCACATCCCACAT
TCTTCTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA
GATGAAGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCTGAGCTGGAGGAAGAAGG
CTGACATTACATTAGTTGCTCTCACTCCATCTGGCTAAGTGTGATCTTGAATACCACCTCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTGAAGTAAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCTAATGGTTGTTCATTATATTACACTTTCAAGTAAAAAA

FIGURE 84

MALMILSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAASFCLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVPPLISITGYVDRDIQLLCQSSGFPRPTAKWKGPGQGDLSTDRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAEELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTTHRKAPQEVPHSEKRF
TRKSVVASQSfqAGKHYWEVDGGHNKRWRGVCRDDVDRRKEYVTLSPDHGYWVRLNGEHLYFT
LNPRFISVFPPRTPTKIGVFLDYECGTISFFNINDQSLIYTLCRFEGLLRPYIEPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGGGCCCTGGCACCTCTAACCCCAGACATGC TGCTGCTGCTGCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCGTGTGTCCATGTGCCCTGCTCCTCTCCTACCCCTCGCATGGCTGGATT
TACCCCTGGCCCAAGTAGTTCATGGCTACTGGTTCGGGAAGGGGCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGTGATGCCGGGAGA
TACTTCTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCACCACCCGCTCCTCGGTGCTCACCCCTATCCCACA
GCCCGAGGACCATGGCACCAAGCCTCACCTGTCAAGGTGACCTTCCCTGGGCCAGCGTGACCA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCTCAGAACCTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTCATCTGTCACTCCAGAGGGCCAGTCT
GCGCCTGGTCTGTGAGTTGATGCAAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCCTGTGCCCTCACAGCCCTCAAACCCGGGGTGTGGAGCTGCCCTGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCCTCGGTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGAUTCAGGGGTGGTCGGGGAGCTGGAG
CCACAGCCCTGGTCTCCTGTGCCCTCACAGCCCTCAAACCCGGGGTGTGGAGCTGCCCTGGTGCAC
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAAGGGTTC
AGCCTCTAGGGGCCCTGACTGAACCTGGCAGAACAGCAGTCCCCCAGACCAGCCTCCCCAG
CTTCTGCCCGCTCTCAGTGGGGAGGGAGGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCAGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTGTAGAATTAAACAGCCCTAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAGAACACAGGCTTGTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTATTTTTAACTAAAAGACAGACAAATTCCA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHPCFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLISRARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSVWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTSLTCQVTFP GASVTTNKTVHLNVSYPPQNLTMVFQGDGTVSTVLGNSSL
SLPEGQSLRLVCADVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAEFTCRAQNP
LGSQQVYLNVSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVERGSASQGPLTEPWAEDSPPDQPPPASARSSVGEHELQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTGTAC
CCAAGGAAAGTCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTCCTGCTGTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTGA
TGGCCTGTATTTCTCCGACTGAGAATGGTGTATCTACCAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGC
CAAACATACAACACCTTGGATCTGCAGAGGCCAGCGATGACTACAAGAACCCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCAATAAGTCCCCCATGCAGCACTGG
AGAAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCAGACACTGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTGGCAGGCCAGAAAACAGCATCTTATTACTCACCCAT
GGCCAGCGGAATTCACTGCAGGGATTGTTCAAGTCAGGGTATTAAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTCTGGTTTGATTGGAGTGGATAT
GGAACATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAACCCAGACCTCTCCTCCAAACCATGAGATCCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGTLVASVHENDMRGKCTVGDRSSQQGSKADYPEGDGNWANYNTFGSAAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIPIVYDFGDAQKTASYYSPYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGC~~GGGGAGACTTCAGGAGTCGCTGTCTGAAC~~TCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCCGAGGGCATGGGCCGGTCTCAGGGCTTGTGCCCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTATTCTGGTCCC~~GGACAGCAACATACAGGC~~CTG
CCTGCCCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATT~~CAGCTGGTGCCGCGCT~~T
CTGTCACCC~~TGGGCTCTTG~~CAGTGGAGCTGGCGGTTCC~~TCA~~GGAGTCTCCATGTTAAC
AGCACCCAGAGCCTCATCTCCATTGGGCTCACTGTAGTGCATCCGTGGCC~~CTG~~C~~TTCTT~~CAT
ATTCGAGCGTTGGAGTGC~~ACTACGTATTGGTACATTTTG~~TCTGCAGTGCC~~CTTCC~~AGCTG
TCACTGAAATGGCTTATT~~CGT~~ACCGT~~CTTGGGCTGAAAAAGAAACCC~~TTCTGATTACCTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCCG~~CTTC~~GTATTCC~~TGG~~AAGAAGGAAG
GCATAGGCTTCGGTTTCCCCTCGGAAACTGCTCTGCTGGAGGATATGTGTTGGAAATAATTACG
TCTTGAGTCTGGGATTATCCGATTGTATTAGTGC~~TTG~~TAA~~AAAATATGTTG~~TAGTAACA
TTAAGACTTATATA~~CAGTTAGGGACAATT~~AAAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFV
TVFGLKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

2020-06-03 10:20:30

FIGURE 91

CTGGGACCCGAAAAGAGAAGGGAGAGCGAGGGACAGAGCGGAGGAAGATGCAACTGAC
TCGCTGCTGCTCGTGTCCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGCCAGGATGATG
GTCCTCCGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCAGCCCCGGCCGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCCCTGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCCTGGGAGGCTTGGGCATTCTTGGCAGCCCCAACCGCCCAGACCACAGCCCC
CACCTCAGCCAAGGTGAAGAAAATCTTGGCTGGGCACTTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTGTACAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCATCAGCCTCGTGCCTCCAGTAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAATCTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGCCGCCGACCTCGCTTGACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACCGCTCAGAGCTCAGGCCACCTGGAGCTGCTCCAGCCCTCAAAGTCGTCTGTG
TCTACATGCCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAAC
CATAGTGATACCCCCTACTACCCATCTGGTGACCCGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAGGACAGGCCCTGCCATGCAGGAGACCATCTGGACACCAGGGCAGGGAAAGGGTTGGCCTC
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGGCAAGTCTCAAGTGGCAG
AGAAAGGGTCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCAGTGG
AGGAGGAGTGGCTCTGTGCAGCCTCACAGGGCTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCCCTGTGGCAGGCCGATCAGTGTGGCCAGATCAAGTCATGGAGGAAGCTAAC
CCTTGGTTCTGCCATCCTGAGGAAGATAGCAACAGGGAGGGGAGATTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCCAGCCCTGCCCTAGTGGCGCCCTGAGCCCCCTGTGCTGTGAGCATGG
CATGAGGCTGAAGTGGCAACCTGGGTCTTGATGTCCTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCCAAAATCCCTCTTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAACAGACAGGACGATTGTGGCTCCACACTAACGGCCACAGCCCATC
CGCGTGCTGTGTGTCCTCTCCACCCCAACCCCTGCTGGCTCCTCTGGAGCATCCATGTCCTG
GAGAGGGTCCCTAACAGTCAGCCTCACCTGTCAGACCGGGTTCTCCGGATCTGGATGGC
CGCCCTCTCAGCAGCGGGCACGGTGGGGGGCCGGCCGCAGAGCATGTGCTGGATCTGTT
TGTGTGCTGTCTGTGGTGGGGAGGGAGGGAGTCTTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGAGCAGGAAATAAGCTGCCCGGGCA

FIGURE 92

MQLTRCCFVFLVQGSILVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVANLLVTGKIVDHGNNTF
SVHFQHNATGQGNISISLVPPSKAVEFHQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPKVVCVYIAFYSTDYRLVQKVCPDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGC~~GGCC~~GTTCTCGGCTGCCCTTCATTGCCCTCGGGCTGCCTGCC
CTTTATGTCTCACCATGCCATCGAGCCGTGCGTATCATTCCCTCATGCCGGAGCTTCCTT
CTGGTTGGTGTCTACTGATTCGTCCCTGTTGGTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGAGCGTTGTCTGTCTATATCCAAGAA
ATGTTCGATTTGCATATTATAACTCTTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTATGCGACTGCTGGCCTATGTTCTGGCTGGGCTTGGAAATCATGA
GTGGAGTATTTCTTGTGAATACCCATCTGACTCCTGGGGCAGGCACAGTGGCATTCAT
GGAGATTCTCCTCAATTCTCCTTATTCAGCTTCCATGACGCTGGTCATTATCTGCTGCATGT
ATTCTGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTAGCCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTATAATCCTGGTGTCAGGGCACCTGGGCATTCTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTGCCAAGACAAGAACCTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCGACACTTCCAAACCGCAGACTACATCTTAGAGGAAGCACAACTGTGCCT
TTTCTGAAAATCCCTTTCTGGTGGATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGP GTVGIHGDSPQFFLYSAFMTLVIIILLHVFWGIVFFDGCEKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCLLCQDKNFLLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGGAGGAGTACGATTCTACTGTTTGCTTAGGATCAC
TCGGTCATTACCACAGCTAAACCTGCTTGGACTCCCTCCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTCCTTTAAGTCTGATAACCATTAAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTACACAACATTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCCATTCTGTTCCCGGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCCAGCAGGCCCTCCAACTCCCAGTGGCACAGATGACGACTTGCAGTGACCACCC
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAAATCAGCAAATGGAATTCA
GTAAGCTTTCAAATTTCAACTAAGCTGCCTCGAATTGGTGATACATGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAAATTCTTAATT
TACCTGAAAATATTCTTGAAATTTCAGAAAATATGTTCTATGTAGAGAATCCAACTTTAAAAAA
CAATAATTCAATGGATAAATCTGTCTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTTGAAAAACTGGAaa
aaa

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQFPSLSLIPLTQM
LTLPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTGCCTGCCACCCAGCCCAGGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTGCTGGCAGCTGGACCAAGGGAGCCAGTCTGGCGCTGGAGGGCTGCTCTGACCATG
GTCCCTGCCCTGGCTGTTGTCTCCGCCCCCAGGCTCTCCCAAGGCCAGGCC
AGAGCTGTCTGTTGAAGTCCAGAAAATGTTCCCTTATACTGACCAAGTGTG
CGCTGCCCGTGGAGGGCTGAAGGCCAGATCGTGTCAAGGGACTCAGGCAAGGCAACTGAG
GCCCATTTGCTATGGATCCAGATCTGGCTCTGCTGGTGAACAGGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACAGCTACAGGTACCCGGAGATGCAGGATGGACATGTCTGGGGTCCAC
AGCCTGTGTTGACACGTGAAGGATGAGAATGACCAGGTGCCCTATTCTCAAGCCATCTAC
AGAGCTCGGCTGAGCCGGGTACCCAGGCCAGGCATCCCCCTCCCTCTCCAGGCTTCAAGCC
GGATGAGGCAGGCACAGCCAAGCTCGGATCTCAGTCCACATCTGAGCCAGGCTCCAGGCCAGC
CTTCCCCAGACATGTCCAGCTGGAGCTGGCTGGCCCTCAGCCCCAAGGGGAGC
ACCAGCCTTGACCAACGGCTGGAGAGGACCTACAGCTGGTACAGGTCAGGACATGGGTGA
CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAGTCTCCATCATAGAGACACTGGGTG
CCCTAGAGCCTATCACCTGGCAGAGAATCTCAAAGTCTTACCCGACCCATGGCCAGGTA
CACTGGAGTGGGGTGTGACTATCACCTGGAGAGGACATCCCCGGGACCCCTGAAGTGA
TGCAGAGGAAACCTACGTGACCAAGAGAGCTGGACAGAGAAGGCCAGGCTGAGTACCTGCTCC
AGGTGCGGCTCAGAATTCCATGGCAGGACTATGCGCCCTCTGGAGCTGACGTGCTGGTG
ATGGATGAGAATGACAACGTGCTTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGAAGACTAGACTGTCAGCAGAGGATGCAGATGCCCGGCTCCCCA
ATTCCACGTTGTATCAGCTCTGAGCCCTGAGCCTGAGGATGGGTAGAGGGAGAGCCTTC
CAGGTGGACCCCCACTCAGGCACTGTGACGCTGGGGTGTCCCACTCCGAGCAGGCCAGAACAT
CTGCTCTGGTGTGGCCATGGACCTGGCAGGGCGCAGAGGGTGGCTTCAGCAGCACGTGTAAG
TCGAAGTCGCACTGGAGGACTATCAATGATCACGGCCCTGAGTTCATCACTTCCAGATTGGGCT
ATAAGCCTCCCTGAGGATGTGGAGCCGGACTCTGGTGGCCATGCTAACAGCATTGATGCTGA
CTCGAGGCCGCTTCCGCTCATGGATTGGCATGGAGAGGGAGACACAGAAGGACTTTG
GCCTGGATTGGAGGCCAGACTCTGGCATGGTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCCAAGTCATGGAGGTGGTGGTGGTGCAGAGTGTGGCGAGCTGGTGGGGCAGGCC
AGGCCCTGGAGCCACGCCACGGTGAAGTGGAGAGGATGATGCCACCCCCAAGTTGG
ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTCCCCAGGGGCTTCTCTGACCATC
CAGGCCCTGGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCATGACTCAGAGGGCTGGCT
CTGCATTGAGAAATTCTCCGGGGAGGTGCACACCAGGCCAGTCCCTGCAAGGGGCCAGCTGGG
ACACATCACCGTGTGGAGGCCAGGATACAGCCCTGACTCTGGCCCTGTGCCCCA
TACCTCTGACACCCCCGCAAGACCATGGCTGATCGTGAAGTGGACCCAGCAAGGACCCGATCT
GGCCAGTGGGACGGTCCCTACAGCTCACCTTGGTCCAACCCCAAGGTGCAACGGGATTGGC
GCCTCCAGACTCTCAATGGTCCCATGCCTACCTCACCTGGCCCTGCATTGGTGGAGGCCAGT
GAACACATAATCCCCGTGGTGGCAGCCACAATGCCAGATGTGGCAGCTCTGGTGTGAGTGA
CGTGTGCGCTGCAACGTGGAGGGCAGTGCATGCGCAAGGTGGCCAGTGAAGGGCATGCCA
CGAAGCTGCGCAGTGGCATCTTGAGGCACCCCTGGTAGCAATAGGAATCTCCTCATCTC
ATTTTACCCACTGGACCATGTCAGGAAGAAGGACCCGATCAACCGAGCACCGTGGCC
GAAGGCAGACTGTCTGATGGCCAGGCACTGAGTGGAGCTGGCCCTGGCTCCATCTGAG
TCCCCCTGGGAGAGGCCAGCACCCAGATCCAGCAGGGGACAGGACAGAGTAGAAGGCCCTCCA
TCTGCCCTGGGGTGGAGGCACCATCACCACATCAGGCCAGTGTGCAAGGCCCTGGAGACCCAACCT
TATGGACTGCCCATGGGAGTGTCTCAAATGTCAGGGTGTGGCCCTGGCTCCATCTGAG
CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFLYLTKLPLPREGAEGQIVLSGDGKAT
EGPFAMDPDSGFLLVTRALDREEQAELYQLQVTLEMQDGHVLWGPQPVLVHKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STS LDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDYAAPPLEHVL
VMDENDNVPICPPRDPTVSIPELSSPGTEVRLSAEDADAPGSPNSHVVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVPLLAGQNI LLLVIA MDLAGAEGGFSSTCEVEVA VTDINDHAPEFITSQIG
PI SL PEDVE PG TLV AMLTA IDA DLEPA FRI MDFA IER GDTE GT FGL DWEPDS GHV RL RLC NLS Y
EAAPSHEVVVVVQSVAKLVGPGPGP GATATVTVLVERVMPPP KLDQESYEASVPISAPAGSFLLT
IQPSDPI SRTL RFSL VNDSEGWL CIEKFSGEVHTAQS LQGA QPGDTY TVL VEA QDTA LT LAPVPS
QYLCTPRQDHGLIVSGPSKDPDI LASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTLALHWVEP
REHIIIPVVVSHNAQM WQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGT LVAIGIFLI
LIFT HWTMSRKKD PDPAD SVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCCTGGAGGAAGCCTAACGGAAACCCAGGCATCCAGCTGCCACGCCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGCTAGGAGACCCACGCTCCCTGGAAGCACCAGCCCTTA
TCTCTCACCTCAAGTCCCCTTCTCAAGAATCCTCTGTTCTTGCCTCTAAAGTCTTGTAC
ATCTAGGACCCAGGCATCTGCTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAATG
TTCTCCTATGTTGGTCTACTATTGCAATTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCAGTGGACCCAGCACAGCCACCAACTCTGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCAATGGGTCA
GCATAGTCACCAACTCTGAGTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCAGTGGGCCAGCACAGTCACCAACT
CTGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCAGTGGGCCAGCACAGCCAC
CAAECTGACTCCAGCACAAACCTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCACCTCCAGTGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGCCAGCAGTGGGCCACCAACTCTGAG
TCCAGCACAGCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCC
AGTGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCAC
CAAECTGAGTCCAGCACAGTGTCCAGTGGGCCAGCAGTGGGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGTCAAGCACAGCCACCAACTCTGAGTCCAGCACAAACCTCCAGTGGGCCAGCAC
GCCACCAACTCTGACTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAAACCTCCAGTGGGCCA
ACACAGCCACCAACTCTGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCCACAAACTTCCCCATAGTCATCTACTGCAGTGTGAGGAAAGCCTGGTGGGTCCCTGGT
GCCGTGGAAATCTTCCATCACCCCTGGTCTCGGTTGTGGGGCTCTTGTGG
TCTTCTCTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTAAACAGCTGTCTACCACCCCT
CATGGCTCAACATGGCCTTGGTCCAGGGCTGGAGGAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCTTAACCTGGTTCTGGAGGAGACCACTGATCATCGATAGCCATGGAGATGAGGGAGGA
ACAGCGGGCCCTGAGCAGCCCCCGGAAGCAAGTGCCGATTCTCAGGAAGGAAGAGACCTGGCA
CCCAAGACCTGGTTCTTCTGAGAAACAGCCTGTCCCTGAGAAACACCTTACAGACACTGGAAAGAGAATACTATAT
TGCTCAATTAGCTAACGAAATAATACATCTCATCTAACACACAGCACAAAGAGAAGCTGTGCTTG
CCCCGGGGTGGGTATCTAGCTGAGATGAACCTGAGTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKM**Q**KGNVLLM**F**GLL**L**H**E**ATNSNETSTSANTGSSVISSGASTATNGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSG**I**ATNSESSTTSSGASTATNSESSTPSS
GASTVTNGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTSNGAGTATNSES
STTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGLVPWEIFLITLVVVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYH**P**HGLNHGLGP*G*PGGNHGAPHRPRWSPNWFWR*R*PVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCCGTACGGGATGAATTAAACGGCGGGTCCGCACGGAGGTGTGACCCCTA
CGGAGCCCCAGCTTGCACGCCACACTCGCGTGCACGGCGTGCCTGCTGTCACAGGTG
GGAGGCTGAACTATCAGGCTGAAAAACAGAGTGGTACTCTCTGGAAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTCTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTGAATCAT
GGTGTCAAGGAAAGGGATTACTTACTGACTCTGTTGGGGAAAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTGGTGTAAAAGT
GATTATAACTGGGGATGCATTGTCCTGGAGAAAGAAGTGTCAATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTGGCCATGCGAGCTGCTGCCTATAT
CTTCATTCAAGGAAATGGAAGGATGACAAGAGCCATTGCAAGACATGATTGATTACTTTGTG
ATATTCAACGAACCACTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTACTTGTGGTAGACCGTCAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAAACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTT
CCCAGGGAAATCCACTTCACGTCACCGGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCACAAACGGTGGGAAGAGAAAGAAGAGGCTGCGTCCCTCTATCAAG
GGGAGAAGAATTATTACCGGACAGAGTGTCAATTCCACCTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTATACGTATTGGACCTGTTGCAATGCTGCTACTCAT
ATATTGTACAGTCTGTTAAGTGGTATTATAATCACCATTGTAATCTTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACCTGCATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAATGAGTAAAGATTATAAGGTTGCCATGTGAAACCTAGAGCATATTG
GAAATGTTCTAACCTTCTAACGCTCAGATGCAATTGCTGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTTGCACCTAACATTGTGGGGAAAATTGCTACAATT
TTAATCTCTGAATGTAATTGACATGTCAGCAGGGAGTGATGGGGTGAATAACTT
GGGCCAGAATATTAAACAAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTMDWM
FLWNCLMRYSYLRLIKEKICLKASLKGVPFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRREGKNLDAVHDITVA
YPHNIQPSEKHLLQGDFPREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCACTGTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTTGGTCTCAGTTCTACGAGCTGGTGTAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCAGGCCCTGGTGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT
TTCTGAGACCACTGAGGGTATGAAAGTGCAGGGTCTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGAAGACTGGAACTCTAACAGCAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGTGAAGGACTCCATTGCAAGGGGGCGTGTCTCTAAGGCTAAAAAACATCACTCCCTCG
ACATCGGCTGTATGGGTGCTGGTCTCAGTCCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGCTCACTCCCTCAATTCCATCGTGGATATGTTGACGGAGGTATCCA
GTTACTCTGCTGTCCCTAGGTGGTCCCCAGCCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTGTCTTCAGACTCCAGAGCAAATGCAGATGGTACAGCCTGTATGATGTGGAGACTCC
ATTATAGTCCAGGAAAATGCTGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGACTCATGA
GGTGGAAATCCAAGGTATTGATAGGAGAGACGTTTCCAGGCCCTCACCTGGGCCCTGGCTCTA
TTTACTCGGGTTACTCTGTGGTGCCTGTGTGGTGTGATGGGATGATAATTGTTCTC
AAATCCAAAGGAAAATCAGCGGAACGGACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCGAAACAGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTACCCGAAGCTCTGCG
TTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCAGGAGGTGCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTTGAGGAGTGTGGGATGACGTAGACAGGGGAAGAACAA
ATGTGACTTTGTCTCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTGTATTTC
ACATTCAATCCCCATTATCAGCCTCCCCCAGCACCCCTCCTACAGCAGTAGGGGTCTTC
GGACTATGAGGTGGGACCATCTCTTCTCAATACAAATGACCAGTCCCTTATTATACCTGC
TGACATGTCAGTTGAAGGCTGTGAGACCCATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATTATCATATGTCAGTGTCTGGGATGAGACAGAGAAGACCCCTGCTAAAGGGC
CCCACACACAGACAGCAGACAGCCAAGGGAGACTGCTCCAGGGCTCCAGCTTCTCT
CCGGAGCCTGCGCACAGAGACTCACGCCCTTACTCTCTTCTAGGGAGCTGAGGTTCTCTGCC
TGAGCCCTGCAAGCAGCAGTACAGCTCCAGATGAGGGGATTGGCCTGACCCCTGTGGGAG
TCAGAAGGCCATGGCTGCCCTGAAGTGGGAGCGGAATAGACTCACATTAGGTTAGTTGTGAAAAA
CTCCATCCAGCTAACAGCAGTCTGAAAGATTAAGAGACAAACGAATGTGAATCATGCTGCAGGTT
ACAGTGATTCCCTGCTCACAGGTGAAGATTAAGAGACAAACGAATGTGAATCATGCTGCAGGTT
TGAGGGACAGTGTGCTAATGATGTGTTTATATTACATTTCACCCATAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTCTATACCAAAATCACCCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTCAAAGAATAAAAGAGGAGGTAGGATTTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAACTACAGGTCCATATCC
CTCATTAAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAATATTAA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGCAGAGTTGGTTAATATTAAAT
ATCAACCAGTGTAAATTCAAGCACATTAATAAGTAAAAAGAAAACCATAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFWQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRAGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFQPPTAKWKGPQGQDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELDWRRKHGQAEILDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGVYVGCRDDVDRGKNNVTLSPNNGYWVRLTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTGGCAATGATGTATCGGCCAGATGTGGTAGGGCTAGGAAAAGAG
TTTGGTGGGAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCCTGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCATTACAAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGGCTCTAACAAATTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTGGAAAATGCATTTATAAATCTCCATTAAGGGAAATTGTCAGTCTCAGGTTATCAAGATTC
AGTCAACAGAACGATGGAGTGTGGCTCATATGCTGTTGATTGAGATTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTTCAACTTGTACATGAAAGCTGCAAGATGCTGTAGGACCCCCCTAAAGTAG
ATCCTCACTCACTAAATTAAAAAAATCAACAAGACAGAACAGACAGCTATCTAAACCATTGCTGCCA
ACACGAAGAAGTAAACTCTAGGTAGCTCAGGATCGTGGTGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGATGGAGTCATCGCTGAGCAACCTTAATTGCCACATGGC
TTGTGAGTGCTGCTCACTGTTACAACATATAAGAACCTGCCAGATGGACTGCTCCTTGGAGTAACA
ATAAAACCTTCGAAAATGAAACGGGTCTCCGGAGAATAATTGTCATGAAAATACAAACACCCATCACA
TGACTATGATATTCCTTGAGAGCTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTGTC
TCCCTGATGCACTCATGAGTTCAACCAGGTAGTGTGATGTTGTGACAGGATTGGAGCACTGAAAAT
GATGGTTACAGTCAAAATCATCTCGACAAGCACAGGTGACTCTCATAGCGCTACAACCTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAACAGATGCAT
GCCAGGGTACTCTGGAGGACACTGGTAGTTCAAGTGCTAGAGATATGGTACCTTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAACCCACAAGCCTGGTGTACTAGAGTTACGGCCTTGCAGGACTG
GATTACTTCAAAAATGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTGTTTTG
GGTGTGGAGGCCATTAGAGATACAGAATTGGAGAAGACTGCAAAACAGCTAGATTGACTGATCTCA
ATAAAACTGTTGCTTGTGATGCTATTTCTTCCCAGCTGTCCTGCACGTAAGCATTGCTTCTGCCA
GATCAACTCTGTCATCTGTGAGCAATAGTGAACATTATGTACATAGAGAAATAGATAATACAATTAC
ATTACAGCCTGTATTCAATTGTTCTAGAAGTTGTCAGAATTGACTGTCAGATAATTGTAAT
GCATATATACAATTGAAAGCACTCCTTCTTCAGTTCTCAGCTCCTCTCATTTCAAGCAAATATCCATT
TCAAGGTGCAAAACAGGAGTGAAGAAAATATAAGAAGAAAAAAATCCCTACATTATTGGCACAGAA
AAGTATTAGGTGTTCTTAGGAAATTAGAAATGATCATATTCAATTGAAAGGTCAAGCAAAGACA
GCAGAAATCCAATCATTTCACTTCAAGGAAAGTGAAGAAAATGTCAGGAAAGTCCAGAAAGCCAAG
ATATATCCTTATTTCATTTCCAAACAACTACTATGATAATTGAGAAAGATTCTGTTTTGACCT
ATAATAATTACAAACTCATGCAATGTAATTGTTCTAACGAAATTAAAGCAATTATTTATTAAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAYKSPLEEfvKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKGSL
RRIIVHEKYKHP SHDYD ISLAELSSPV PYTNAVHRVCLPDAS YEFQPGDVMFVTGF GALKNDGYS
QNHLRQAQVTLIDATT CNEPQAYNDAITPRMLCAGSLEGKTDACQGD SGGPLVSSDARDI WYLAG
IVSWGDEC A KPNK PGVY TRV TAL RDWITS KTG I

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAACGCTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAACGTTCCCTG
CCCCGATGAGCCCCCGCCGTGGTCCCCGACTATCCCAGGGGGCGTGGGCACCGGGCCAGC
GCCGACGATCGTGCCTGGGAGTAGGATGTGGTGAAGGATGGGCTTCCTT
ACGGGGCTCACAATGGCCAGAGAACGATTCCGTGAATGTCGCTGCCTGCTACGCCCTCAA
TCTGCTCTTGTTGTAATGTCCCATCAGTGTGTTGCACTTCTGCTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTAACAGAACGAGGTAGAGGAAGCAGTCATTTGACTTACTTCCT
GTGGTTCATCCGGTCACTGATTGCTGTTGCTGTTCTTATCATTGTTGGGATGTTAGGATATTG
TGGAACGGTGAAGAACGATCTGTTGCTTGCATGGTACTTTGGAAGGTTGCTTGTCACTTCT
GTGAGAACGGTGTGGCTTGGACATATGAACAGGAACCTATGGTTCCAGTACAATGGTCA
GATATGGTCACTTGAAGGCCAGGATGACAATTATGATTACTAGATATGGCTTACTCA
TGCTTGAATTTCAGAGAGGTTAAGTGTGGAGTAGTATATTCACTGACTGGTGG
AAATGACAGAGATGGACTGGCCCCCAGATTCTGCTGTGTTAGAGAATTCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAGGGTGTGGAGAACGTT
TTTGAGAGGAACCAACACTGCGAGGTGCTGAGGTTCTGGAAATCTCATTGGGTGACACAAA
TCCTGGCCATGATTCTCACCATTAACCTCTGCTCTGGCTCTGTATTATGATAGAAGGGAGCCTGGG
ACAGACCAATGATGTCTTGAAGAATGACAACACTCTCAGCACCTGCTGCTCCTCAGTAGAACT
GTTGAAACCAAGCTGTCAAGAATCTTGAACACACATCCATGGCAAACAGCTTAATACACACT
TTGAGATGGAGGAGTTAAAAAAGAACGTTAAGAAGAAAACCACAAACTTGGTTATTGGACT
TGTGAATTGGAGTACATACTATGTGTTAGAAATATGAGAAATAAAATGTTGCCATAAAA
TAACACCTAACGATATAACTATTCTATGCTTAAAGGAGTGGAAAAGTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAAGCTGCTGAAGACAGATGTCATACCAACTGTGAGCC
TGTGTATGACTTTACTGAACACAGTTATGTTTGAGGCAGCATGGTTGATTAGCATTCCGCA
TCCATGCAAACGAGTCACATATGGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
CTAGTATATAAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
CTCAGCGATCTATTCTCTGATGCTAACATTTATATCAGAAAACCTTCAATATTGGTACT
ACCTAAATGTGATTGGTCTGGTTACTAAATATTCTTACCACTTAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTGGTATATAAAGCTTAAATCTGTATAATTCACTGCGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTGAAAAGAAAATTGTCCTGTATAGCATCATT
ATTGGTCTTCTGGCTTCTGTTAAATAAGCTTACTATTCTGCTGGCTTATATACACATATAAC
TGTTATTAAATACCTAACACTAATTGAAAATTACAGTGTGATACATAGGAATCATTATTC
AGAATGAGTCTGGCTTAGGAAGTATTAAAGAAAATTGACACATAACTTAGTTGATTGAGA
AAGGACTTGTATGCTGTTCTCCAAATGAAGACTTTTGACACTAACACTTTTAAAAAA
GCTTATCTTGCCTCTCCAAACAGAAGCAATAGTCTCAAGTCAATATAAAATTCTACAGAAAA
TAGTGTCTTCTCCAGAAAATGCTGTGAGAATCTTAAACATGTGACAATTAGAGATT
CTTGTGTTATTCACTGATTAATATACTGTGGCAATTACACAGATTATTAATTTTACAA
GAGTATGTTATTGAAATGGAAAAGTGCATTACTGTGTTGTGTTGTTATTTGTTAT
TTCTCAGAATATGAAAGAAAATTAAATGTGCAATAATATTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNVLTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMNTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMWDPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRLFGISIGVTQILAMILTITLLWALYYDRREPDTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGGACACCTTATCCCACTCATCCTCATCCTCTTCTGATAAAGCCCCTACCAGTGCT
GATAAAAGCTTTCTCGAGAGCCTAGAGGCCTAAAAAAAAGTGCTGAAAGAGAAGGGGACAAGGAACA
CCAGTATTAAGAGGATTTCCAGTGTTCAGTGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTTCTCACCTG
CCTCTCATCACAGGCACCTCGTGCACCCGTGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAAGTGGATGAGTCACAGTCCTCTATGTGACAACCAGTGAATGGGAG
TGGTACCACTTCACGGCATGGGGAGATGCCATGCCTACCTTCTGCATACAGAAAACACTGTGGAACCC
CGCACCTGCTGGCTCAATGGCAGCCACCCCTAGAAGGGCAGGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
TCAATGGAACTGCTCTCTGGAACACCACGGTGGAAAGTCAGGCTTGCCTGGAGGCTACTATGTATCGT
CTGACCAAGCCCAGCGCTGCTTCCACGTCTACTGTGGTCACTTGTGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATAACCGAGTCACATGCGCTCCAGGAACGTGCTAGGCCCTGACAGGCAGACATGCTTGT
ATGAAAATGAATGTGAGCAAACACGGTGGCTGAGTGTGAAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTTGGCGTGTGCTAAGAAGTGTGCAAGACTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTGCCTGGATCTGAGAAAGGCTACAGTGTGAATGTCCCCGGGCCTGGTGT
CTGAGGATAACCACACTGCCAAGTCCCTGTGTGCAAATCAAATGCCATTGAAGTGAACATCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTTCTGACCAACACCTCCCTGCCAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTCACTGTGGAATGACAAGATTGTGGCAGCAACCTCGTGA
CAGGTCTACCAAGCAGACCCGGGGAGCAGCGGGACTTCATCATCCGAACCAAGCAAGTGTGATCCCAGGT
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACGTTCCAACTTCGAAACTCCCCACTGAAAT
CATGAGCCGAAATCATGGATCTCCCATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTAC
GGGAAGCTCTGCCACCCCTCAAGCTCGTGAETCCCTACTTGGCATTGAGCCGTGGTGCACGTGAGCGGC
TTGGAAGCTTGGTGGAGAGCTGCTTGCACCCACCTCCAAAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAAGATGACTCGTAAAGCAGTACACATCCCGGATCACCTAGCAAAGCATTCCAGG
TCCCTGCTTCAGGTTGTGGCAAGACCACAAGGAAGTGTCTGCACTGCCGGTTCTGTCTGTGGAGTG
TTGGACAGCGTCCCGTGTGCCACGGCAATGGCTGTGGAGGACTCAGGTCAGGAGGACTCAGC
CGGTCTACAGGGCCAGCCTAACAGCGGCCGATCCGCATCGACTGGAGGACTTAGTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTTGGAGCTCTCCCCCACCCTCTAAGAACATCTGCCAACAGC
TGGGTTAGACTTCACACTGTGAGTTCAAGACTCCCAGCACCACACTCTGATTCTGGTCATTCACTGGCA
CAGGTACAGCACTGCTGAACAATGTGGCTGGGTGGGTTCATCTTCTAGGGTTGAAAACACTAAACTGTCCA
CCCAGAAAGACACTCACCCATTCCCTCATTTCTTACACTTAAACACTCGTGTATGGTGCACAGAC
CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCTAGAAAAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTTAAATATGTAATTAGTTACCTTGAAATTCAATTCAAATGCAGACTAA
TTATAGGAATTGGAAGTGTATCAATAAAACAGTATATAATT

FIGURE 110

MPPFLLTCLFITGTSVSPVALDPCSAISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPEHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCLWNNTVEVKACPGGYYVYRLTKPSVCFH
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGCASHCLGSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFILTNTSCRGVSNGTHVNLFLSLKTCGTVVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPV
CEFPPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFITLEIFKDNEFEEPYREALPTLKLRLDSLYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYILIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAGCHRRMRRGAGGEDSAGLQQQLTGGLP RIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCTGCACACTGG
GCCTCCCTCCAGCAGTGTGACCAAGGGACTTCAGCTGACCTGCTGGCCAGCAGGACCTGTGTGGGAGGCCCT
CCTGCTGCCCTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGTACAAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGACGACTGGCGAGTATCATCATTGT
GGTTGCTCTCATCAAGGTGATTCTGGATAAAATACTACTTCCTCTGCAGGCTCTCCACTTCATCCGA
GGAAGCAGCTGTGACGGAGAGCTGGACTGTCCTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCCGCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAACTGGTCTCTGCCTGTTGACAACTTCACAGAAGCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAGGATCTGGATGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATCGGAACACTCAAGTGGCCCTGTCCTCAGGCTCCCTGGTCTCCCTGCACTGTCTGCCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGGAGGAGGAGGCTCTGGATTCTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGTCCTCACGGCA
GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGCAG
CTTCCCCTCCATGGCTGTGGCCAAGATCATCATCATTAACCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTCCACTCACTTCTCAGGCACAGTCAGGCCATCTGTCTGCCCTTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATGGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGTCGACATACTGCTGCAGGCGTAGTCCAGGTCTGGACAGCACACGGTGCAATGCAGACGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGTCAGGCATCCCGGAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGATGTACCAATCTGACCAAGTGGCATGTGGTGGCATCGTTAGCTGGGCTATGG
CTGCGGGGGCCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCTTGCAAGTGTGGAGGCGCTTCCCTGCCCTGCCACCT
GGGATCCCCAAAGTCAGACACAGACAGAGGTCCCTGGGTACACCCCTGCCACAGCCTCAGCAT
TTCTGGAGCAGAAAGGGCTCAATTCTGTAAGAGACCCCTGCAGCCAGAGGCGCCAGAGGAAGTCA
GCAGCCTAGCTGCCACACTGGTGTCCAGCATCCAGGGAGAGACACAGCCACTGAACAAGGTCT
CAGGGGTATTGCTAACGCAAGAAGGAACCTTCCACACTACTGAATGGAAGCAGGCTGTCTTGAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTCTGCCAGCCCTGTCCGTCTCACCCATCCCCAA
GCCTACTAGAGCAAGAACCAAGCTGTAAATAAAAATGCACGCCACTGTGGTATGACTACCGTTACCT
ACTGTTGTCAATTGTTATTACAGCTATGCCACTATTATTAAAGAGCTGTGTAACATCTGGCAAAAAAAA
AAAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIVVVLIKVILDKYFLCG
QPLHFI PRKQLCDGELDCPLGEDEEHCVKS FPEGPAVAVRLSKDRSTLQVLD SATGNWFSACFDN
FTEAL AETACRQM GYSRAVEIGPDQ DLDV VEITEN SQEL RMRN SSGP CLSGS LVSL HCLACG KSL
KTPR VVGEEASV DSWPWQVS I QYDKQHVC CGGS ILDPH WLTA AHCF RKHT DVFNWK VRAG SDKL
GSF PSLA VAK IIIIE FNPM YPKDN DIA LMKL QFPL TFS GTVR PICL PFF DEEL TPAT PLWI IGWG
FTKQ NGG KMS DILL QASV QVID STRC NADDAY QGEV TEK MM CAGI PEGG VDT CQGD SGGL MYQS
DQWHVVGIVSWG YGC GGP STPGV YT KVSAYLN WIYN VWK AEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACTCCTGGTCCAAGTGATCCACCCGCCTCAGCCTCCAAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTGGCCTCTGAACAACCTTTCACTGCAACTAAAAAGCCACAGGAGT
TGAACGTAGGATTCTGACTATGCTGTGGCTAGTGCTCTACTCCTACCTACATTAAAATC
TGTTTTTGTCTCTGTAACTAGCCTTACCTTCAACACAGAGGATCTGTCACTGTGGCT
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGTCCCCTCGAAG
CCGGGGACAGCCTCACCTGCTGGCCTCGCTGGAGCAGTGCCTCACCACACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTTGGTAGCTGCGGCTTCAGGTGGC
CTTGCCCTGGCCGTAGAAGGGATTTGACAAGCCCCGAAGATTTCATAGGCGATGGCTCCACTGCC
AGGCATCAGCCTTGCTGTAGTCATCACTGCCCCGGCCAGGACGGGGCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTTCATGTCCTGCACATCACCTG
ATCCATGGGCTAATCTGAACCTGTCCAAAGGAACCCAGAGCTTGAGTGAGCTGCGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAACTGTGCCAATTATGGGTAGAAAAGATG
GAGGTGTTGGGTTATCACAGGCATCGAGTCTCCTGCATTCACTGAGCTGACATGTGGGGAAAGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCGTTCCGCCCGAT
CCACGTACCGAGTGTGAAGGGCAACTGCAGGCCATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGCAGCCGTGGGAAGGAGCAAGCAAAGTGACCATTCTCCCTCCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTAGTGTGGCCAGGAGGTAGCAAGGCCTGAGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAATCCCAGGCAAAGGACTGTGTGG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCAAAGGGCCCTATACCCCAGGAGACTTGTGATTTGAAACCCCAAATCCA
AACCTAAGAACCAAGGTGCATTAAGAACATCAGTTATTGCCGGTGTGGCTGTAATGCCAACAT
TTTGGGAGGCGAGGCGGGTAGATCACCTGAGGTAGGCTAAGGACAGCCCTGGCCAACATGG
TGAAACCCCTGTCTACTAAAAACAAAAAAACTAGCCAGGCAAGGAGCTGGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTGAAACCTGGAGGTGAAGGAGGCTGAGACA
GGAGAACATTCAGCCTGAGCAACACAGCGAGACTCTGTCAGAAAAAAATAAAAAAGAATTA
TGGTTATTTGAA

FIGURE 114

MLWWLVLLLLPTLKVFCSLVTSLYLPNTEDSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLAELTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGCTCTCAGCCTCTCAAAGCAAGGAAAGAGACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAATCACTTAAGATTGTGGACTGGTGTGTTGCTGCCCTAACTCTAATTGTCCT
GTTTGGGGAGCAAGCACCTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAACCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATAACTGG
CATCTACTCGTGGGTCTCAAAAATGTTTATCAAACAGATAAAGTGATTCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTCTTGAAACAGTCAGTG
ATTGGGTCCCAGCAGAAAAGCCTATTGAAACCGAGATTCTTAAAAATTCCAAAATTCTGG
GATTGAGTACGTGACCATGTATTGGATCAATCCACTCTAATATCAGTTCTGAGTTACAAG
ACTTGAGGAGGGAGGAGAAGATCTTCACTTCCCTGCCAACGAAAAAAAAGGGATTGAACAAAAT
GAACAGTGGTGGTCCCTCAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTCCAATAATGACTATACTGAAAATGGAATAGAATTGATCCCATGCTGGATGAGAG
GTTATTGTTGATTACTGCCGTCGAGGCAACCGCTATTGCCGCCGTCTGTGAACCTTACTA
GGCTACTACCCATATCCACTGCTACCAAGGAGGACGAGTCATCTGCGTGTACATGCCCTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTAATAGGAGGTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAATGCATGCTATTCAATGAATTCTGCCATGAGGCATCTGGCCCT
GGTAGCCAGCTCCAGAATTACTGTAGGTAATTCTCTCTTCAIGTTCTAATAAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKKICKSLKICGLVFGILALTLIVLFWGSKHFVPEVPKKAYDME
HTFYNSGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEETTTFFEQSVIWPVPAEKPIENRDFLKN SKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMED
ERGYCCYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTTCACACCTCGGCAGCAGGAGGGCGCAGCTCTCGCAGGGCA
GGCGGGCGGCCAGGATCATGTCACCACCATGCCAAGTGGTGGCGTCCCTGTCATCCTGGGCT
GGCGGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCC
TATTTCACCATCTGGACTTCCAGCATGCTGCAGGCAGTGCAGGCCGTACGATCGTAGGCATCGCCT
GGGTGCCATTGGCCTCTGGTATCCATCTTGCCTGAAATGCATCCGATGGCAGCAGTGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGATCATGTTCATGTCTCAGGTCTTGTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAACTCTGGATGTCACAGCTAACATGTACACCGGCATGGTGG
GATGGTGCAGACTGTTCAGACCAGGTACACATTGGTGCAGCTCTGGTGGCTGGCTGGTGCCTGGAGGCC
TCACACTAATTGGGGTGTATGATGTGCATGCCCTGCCGGGCTGGCACCAGAAGAAACCAACTACAAA
GCCGTTCTTATCATGCCCTAGGCCACAGTGTGCCTACAGCCTGGAGGCTCAAGGCCAGCACTGGCTT
TGGGTCCAACACAAAAACAAGAAGATATACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGACTATGTGTAATGCTCAAGACCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCCTAGATTTCTTCTGTTGACTCACAGCTGGAGTTAGAAAAGCCT
CGATTCATCTTGGAGAGGCCAAATGGCTTAGCCTCAGTCTGTCTCTAAATATTCCACCATAAAACA
GCTGAGTTATTTATGAAATTAGAGGCTATAGCTCACATTTCAATCCTCTATTCTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTAATCTCTCTCACATTGATGATTAGACAGACTCCCCCTC
TTCCCTCTAGTCATAAAACCCATTGATGATCTATTCCCAGCTTACCCAGAAAACCTTGTAAAGGAAA
GAGTAGACCCAAAGATGTTATTTCTGCTGTTGAATTGTCTCTCCACCCCAACTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTGTAATCTCTCCAGCCCATGATCTGGTTCTT
ACACTGTGATCTAAAGTACCAAACCAAGTCATTTCAGTTGAGGCAACCAACCTTCTACTGCTG
TTGACATCTCTTATTACAGCAACACCATTCTAGGAGTTCTGAGCTCTCCACTGGAGTCCTCTGT
CGCGGGTCAGAAATTGCTCTAGATGAATGAGAAAATTATTTTTAAATTAACTGCTAAATATAACTAA
ATAAAATAATGTTTAGAAATGATACACTATCTCTGTGAAATAGCCTACCCCTACATGTGGATAGAAG
GAAATGAAAAAAATAATTGCTTGAATTGTCTATATGGTACTTTGTAAGTCATGCTTAAGTACAATTCC
ATGAAAAGCTCACACCTGTAATCTTACCACTTGGAGGCTGAGGAGGAAGGATCACTTGAGGCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTACAAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATACACCTGTAATCTTACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
TGGGGCTGCAGTGAGGCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
ATAAAAAAAATAATGAAACACAGCAAGTCCTAGGAAGTAGGTTAAAACTAATTCTTAA

FIGURE 118

MSTTCQVVAFLSILGLAGCIAATGMDWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHHSVAYKPGGFKA^VTGFGSNTKNKKIYDGGARTEDEVQSYP SKHDY

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTTCTGTGGCACAGAGAACCTGCTCAAAGCAGAAGTAGCAGTCCGGAGTCC
AGCTGGCTAAACATCCCAGAGGATAATGGCAACCCATGCCTAGAAATCGCTGGCTGTTCTTG
GTGGTGGATGGGGCACAGTGGCTGTCAGTGCCTCAGTGGAGAGTGTGGCATGCTTCAATT
GAAAACAACATCGTGGTTTGAAAACCTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGGCCAGAG
GACTGATGTTGCTGCTTCCGTATGTCCTCTGGCTTCATGATGCCATCCTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGTCAGGCTGGAATCATCTCAT
CATCACGGCATGGGGCTCATCCCTGTGAGCTGGGTGCAATGCCATCATCAGAGATTCTATA
ACTCAATAGTGAATGTTGCCAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGCCTGATTGTTGGAGGAGCTGTTCTGCGTTTGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTCCCATCGCACACCCAAAAAGTTATCACACCGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTTAGTTGTTATGTTTTAACTTACTATAAGCCATGCAAATG
ACAAAAATCTATATTACTTCTAAAATGGACCCCCAAGAAACTTGTATTACTGTTACTGTAATTCTCC
AATCTTAATTACAGGAACGTGCATCAGCTATTGATTCTATAAGCTATTCAGCAGAATGAGATA
TTAAACCAATGTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGGTCAAGCATCTA
CTCTTTTATCATTACTCAAAATGACATTGCTAAAGACTGCATTTTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTTATGGT
TTTATTAAATGAAATGCCAGTCCATTACACTGAATAAAATAGAAACTCAACTATTGCTTTCAAGGGAA
ATCATGGATAGGGTGAAGAAGGTTACTATTAAATTGTTAAAACAGCTTAGGGATTATGCTCCA
TTTATAATGAAGATTAATGAAGGCTTAATCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTATCCTCTCCAGAGGCTTTTT
CTTGTGTTAAATTAAACATTAAAACGGAGATATTGTCAAGGGCTTGCATTCAAATGCTT
TTCCAGGGCTACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGATGGTTAGGAAAGTG
AAAATTTTGTGTTGTATTGAAGAAGAATGATGCAATTGACAAGAAATCATATATGTATGGAT
ATATTAAATAAGTATTGAGTACAGACTTGAGGTTCATCAATATAAAATAAGAGCAGAAAATA
TGTCTGGTTTACCTGCTACCAAAAAACACAAACAAAAAAAGTTGTCCTTGAGAAACTCACCT
GCTCCTATGTGGGTACCTGAGTCACATTGTCATTGTTCTGTGAAAATAATTCTTCTGTA
CCATTCTGTTAGTTACTAAAATGTAAACTGTATTGTTCTGTGAAATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTCGACGTCTGTCTAGCTAAATGAATGTGTTCTATTGCTT
TATACATTATTAATAAAATTGTACATTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPQWRVSIFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSHTGKKSPSVYSRSQYY

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAGGCGCATTGATGCAGCCTGCGGCGGCCTGGAGCGCGGGAG
CCAGACGCTGACCACGTTCCCTCCTCGGTCTCCCTCCGCCTCCAGCTCCGCGCTGCCCGCAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCCCTCCCCGAGCGGCTCCGCGGCCTCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCTGAGCCCTGTAGATCCCCAAGGGAAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGCAATGTTATTCCGGTACACCTGGGATCCAGGTGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTGAGGAGTCCTGGACACCCAAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGTCTAACAGAGTTTGGTCAGTGGCTCACTCGGCTAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTCACATTCAATGGAGCTGAATGTTCAGGACCTCTCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGGCCCTGAAATGAATTCAACAATTAAATTATCG
CACTTCTCTGTGGAAGGACTTGTGAAGGAATTGGTGTGGATTAGTGGATGTTGCTATCTGG
TTGGCACTGTTCAAGTACCCAAAAGGAGATGCTTCACTGGATGGAATTCAAGTTCTCGCATC
ATTATTGAAGAACTACCAAATAAATGCTTAATTTCATTTGCTACCTCTTTTTATTATGCC
TTGGAATGGTCACTTAAATGACATTAAATAAGTTATGTATACTCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAAGTGTGATTCACACTGTTAAATCTAGCATTATTCAATTG
CTTCAATCAAAAGGGTTCAATATTTTTAGTTGGTTAGAATACTTCTTCATAGTCACATT
CTCTCAACCTATAATTGGAAATTGTTGTGGCTTTGTTCTCTTAGTATAGCATT
AAAAAAATATAAAAGCTACCAATTGTACAATTGTAAGAATTGGGATATCTGT
TAAATAAAAATTATTCCAACA

FIGURE 122

MRPQGPAAASPQLRGLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLOGPAGVPGR
DGSPGANVIPGTPGIPGRDGFGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGS1RLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDAIWVGTCSDYPKGDASTGWNSVSRIIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGGCGGTACGGGGCTCTCCTGCCTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGAACTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATACATCATGTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCCAAAAAACTGTAAGATGCAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGAATTGATGTTGCTGCGAATGCGGTGTGGGATTATTGTTCTGGAG
TGTTCCTGCTGGCTGGCAAAGAATAATGTCCAAAATCGTCCATCTCCAAGGGTCCAATT
TCTTCTGGGTGTCAGCAGCCTGACTCACTACAGTGCAGCTGACAGGGCTGTCACTGCAACTG
GCCCTAAGCAAAGCAAAGACCTAAGGACGACCTTGAACAATACAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTAGCCCCACTGTCTTACTGACAATG
CTTTCTCTGCCGAACGAGGATGCCCTAACGGCTGTAGGTGTGAAGGCAAATGGTATATTGTGA
ATCTCAGAAAATTACAGGAGATAACCTCAAGTATATCTGCTGGTTAGGTTGTCCCTCGCT
ATAACAGCCTCAAAAACCTTAAGTATACTAATTAAAGGGCTAACCGCTCACCTGGCTATAC
CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTAATGGAATACGCAACTCAAAGA
GCTGATTCTTAGTCCAATAGAATCTCCTATTCTAACAAATACCTCAGACCTGTGACAAATT
TACGGAACCTGGATCTGCTCTATAATCAGCTGCAATTCTGGGATCTGAACAGTTGGGGCTTG
CGGAAGCTGCTGAGTTACATTACGGTCTAACCTCCCTGAGAACCATCCCTGTGCGAATATCCA
AGACTGCCAACCTGGAACCTGGACTTTGGACCTGGGATATAACCGGATCCGAAGTTAGCCAGGAATG
TCTTGCTGGCATGACTCAAGAAACTCACCTGGAGACATCAATTTCAGCTCAACCTGGAT
CTGGCCCTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTACTTGCAAGTGGAAATAAAATCAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTAACAAAGGCTGATTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGTCCTGAATCTGCAGCGCTCAACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTGGATTCTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATGGGAATGCAGCAGAAAATTGCTCCCTGTAAACTGGCTGAAAAGTT
TTAAAGGTCTAACGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCAGTGAAGAAACTACAGCATCTGTCAGCTGGCAAAAGTACTACAGAGAGGTTGATCTGGCCAG
GGCTCTCCCAAAGCCGACGTTAACGCCAGCTCCCAGGCCGAAGCATGAGAGCAAACCCCTT
TGCCCCCGACGGTGGGAGCCACAGAGCCCCAGAGACCGATGCTGACGCCAGCACATCTCT
TTCCATAAAATCATCGGGGCGACGCTGGCGCTTTCCGTGTCCTGTCATCCCTGCTGGTTAT
CTACGTGTCATGGAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAAGCGCTCCCTCATGCGAA
GGCACAGGAAAAAGAACAGACTCCCTAACAGCAAATGACTCCCAGCACCAGGAATTATGTA
GATTATAAACCCACCAACACGGAGACCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATTGAACCATTTGTGATAAAAAGAGCTTAAAGCT
GGGAAATAAGTGGTGTATTGAACCTGGTGAATCAAGGGAACGCGATGCCCTCTCCCC
TTCCCTCTCCCTCACTTTGGTGGCAAGATCCTTCTGTCCGTTTAGTGCATTATAACT
GGTCATTTCTCTCATACATAATCAACCCATTGAAATTAAATACCAATCAATGTGAAGCTT
GAACTCCGGTTAATATAACCTATTGTATAAGACCCCTTACTGATTCCATTAAATGTCGCAATT
GTTTAAGATAAAACTCTTCATAGGTAAAAAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGRCERGMVYCESQKLQEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIENAFNGIRRLKELILSSNRISYFLNNNTFR
PVTNLRLNLDLSYNQLHSLGSEQFRGLRKLLSLHRSNSLRTIPVRIFQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKELHLEHNQFSKLNLAFFPRLVSLQONLYLQWNKISVIGQTMSWTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNSNKLFIGQEILDWSISLNIDSLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLIVILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKKRQSLKQMTPSTQEFYVDYKPTNTETSEMLLNGTGPCYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATGTCGTCCGGAGGAGGAGAGGCTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCTACTGTCGGCTGCCGGCTACCGTGGCCGAGCT
AGCAACCTTCCCTGGATCTCACAAAAACTCGACTCAAATGCAAGGAGAAGCAGCTTGCTC
GGTGGGAGACGGTCAAGAGAACTGCCCCATAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTACAGACACGTAGT
GTATTCTGGAGGTCAATGGTACATATGAACATCTCGAGAGGTTGTGTTGCCAAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTATTGCCAGTT
TTAGCCAATCCAAC TGACCTAGTGAAGGTTAGATGCAAATGGAAGGAAAAGGAAACTGGAAGG
AAAACCATTGCGATTCGTGGTGTACATCATGCATTGCAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTGGGCAAGGCTGGTACCAATATAACAAAGAGCAGCAGTGGTAATATGGGAGATTAA
ACCACTTATGATACTGAAACACTACTTGGTATTGAATAACACCACCTGAGGACAATATCATGAC
TCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTCTATTCTGGAACACCAGCCATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAGGAAGGGACTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGTGAAGGATTCACTGAGTCTATATAAAGGCTTTTAC
ATCTGGCTGAGAATGACCCCTGGTCAATGGTCTGGCTTACTTATGAAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRI MNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTGGCGGCCGGCAGGAGAGCGGCCGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCAGCGAGGCGACGGCGGCATGGCCAGGCCGGCATGG
AGCGGTGGCGCAGCGCTGGCGCTGGTGAACGGGGCTCGGGGCACTGGCGCCGTGGCC
CGGGCCCTGGTCAGCAGGGACTGAAGGTGGTGGCTGCGCCCGCACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAAGGCTACCCCAGGACTTTGATCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCGTTCTCAGCACAGCGGTGTAGACATC
TGCATACAACATGCTGGCTTGGCCGGCTGACACCCCTGCTCTCAGGCAGCACAGGTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGT
CCCCTGTCGTGACCCACTTCTATAGTGCACCAAGTATGCCGTACTGCCGTGACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTGCCCTCAAACCTCACGACAAGGACCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTTTATCTACGTCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGA
TCCTTCCCTCCCCACCCCTCATGGCTGCCCTGCCCTGGATTAGGTGTTGATTCTGGAT
CACGGGATAACCACTCCTGTCCACACCCCGACCAGGGCTAGAAAATTGTTGAGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAATGTGAAAAATGGGCTGGGAAAGGAGGTGGTGT
TAATTGTTTACTTGTAACTTGTCTTGCCCTGGCACTTGGCTTGTCTGCTCTCAGTG
TCTTCCCTTGACATGGAAAGGAGTTGGCCAAACCCATCTTGTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCACCTTATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCCTCTGCCCTGCCCTGCACCCCTCTCCCCCTATCTATCTCCTCTCGGCTCCCC
AGCCCCAGTCTGGCTTGTCCCCCTCTGGGTGATCCCTCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCCTGCCAGTGGATTGATGGTCAATTAAAAAGAAAAATCGAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDILSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLPLSVTHFYSATKYAVTALTEGLRQEELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCTACACCAT
CATGTCCCTCCCACCCCTCTTGACTGCGGGCGTTCAAGGTGCAGAGTCTCAGTTGCCGGGAGC
ACCTCCCTCCCAGGGCAGTCTGCTCAGAGGGCCTGCCAGAATTCCAGTTCTGGTTCATGC
CAGCCTGAAAAGGCCATGGAACTTGGTGAATCACCGATGCCATTAAAGAGGGTTTCTGCCA
GGATGGAATGTTAGGTCGTTGTCTGCGCTGTTCAATTCAAGTAGCCACCAGCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAAATTCTCATGTATTTCATTTATTAA
TTAATTAACTGATAGTTGACATATTGGGGTACATGTGATATTGGATACTGTATAACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTATTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCTCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGGTCAAGCGATTCTCATGCCCTCACCTCCAAGTAGCTGGACTACAGGCAT
GCACCCACAATGCCCAACTAATTGGTATTAGTAGAGACGGGGTTTGCCATGTTGCCAGG
CTGGCCTGAACTCCTGGCCTCAAACAATCCACTTGCCCTGGCCTCCAAAGTGTATGATTACA
GGCGTGAGCCACCGTGCCTGGCTAAACATTATCTTTCTTGTTGGAACTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGACTTCCCTCT
ATCTAACTGTATATTGTAACCAACCGTACTTCATCCCCACTCCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCACCTCCATGAGATCCACTTTTAGCTCCACATGTG
AGTAAGAAAATGCAATTGGTCTTCTGCTGGCTTATTCACCTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTTCGTTCTAATTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSSLRGPRPRI PVLVSCQPV
KGHGTLGESPMPFKRVFCQDGTVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTGTATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCCATCGGACTAAAATTGGCTGTTCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGCGCTGCGATGCCGGTTTCATTTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACCTCTACCTCAGAACACCAAATAATGCTGGATTCTTCAGAT
TTGAAAAACTTGTGAAAGTAGAAAGAATATAACCTATAACCAACAGTTAGATGAATTCTACCAACCT
CCCAAAGTATGTAAGGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCCTATCTGGAAGAATTACATTAGATGACAACCTGTCTGCACTGATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTTCCGTCCGTAATCACCTAGCACAATTCCCTGGGTTT
GCCCAAGGACTATAGAAGAACTACCGCTGGATGATAATCGCATATCCACTATTCATCACCACATCTTCAG
GTCTCACTAGTCTAAAACGCCTGGTCTAGATGAAACCTGTTGAACAATCATGGTTAGGTGACAAGTT
TTCTTCACCTAGTTAATTGACAGAGCTGTCCTGGTGCAGAATTCCCTGACTGCTGCACCAGTAACCT
TCCAGGGCACAAACCTGAGGAAGCTTATCTCAAGATAACCACATCAATGGGTGCCCAAAATGCTTTT
CTTATCTAAGGCAGCTATCGACTGGATATGTCATAAACTTAAGTAATTACCTCAGGGTATCTTT
GATGATTGGACAATATAACACAATGATTCTCGCAACAATCCCTGGTATTGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAA
AGGTTGGTGGGATGGCTATTAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGGGATTGTAAGC
ACCATTCAAGATAACCACGTGCAATACCAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCAGTGAC
CAAACAGCCAGATATTAAGAACCCCAAGCTCAACTAGGATCAACAAACACAGGGAGTCCCTCAAGAAAAAA
CAATTACAATTACTGTGAAGTCTGTCACCTCTGATACCATTCAATCTCTGGAAACTGCTCTACCTATG
ACTGCTTGAGACTCAGCTGGCTAAACTGGGCCATAGCCCGCATTTGGATCTATAACAGAAACAATTGT
AACAGGGGAAACGCACTGAGTACTGGTCACAGCCCTGGAGCCTGATTCAACCTATAAGTATGCTGGTTC
CCATGGAAACAGCAACCTCTACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCCTACAACCACCCCTCAATCGAGAGCAAGAGAAAAGAACCTTACAAAACCCCAATTAC
TTGGCTGCCATCTGGTGGGGCTGTGGCCCTGGTTACCATGGCTTCTGCTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTTTCTCAAGGAACGTGCAATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAAGCTGCCACTAAGAAGGACAACCTATCCTGGAAATCAGGGAAACTCTTCAAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGGAGGAGTTGTAATACACACCATATTCCCTCTAATGGAATGAATCTGT
ACAAAAAAACAATCACAGTGAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTGGGTTTTAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLDNSVSASVIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSILKRLVLDGNLLNNHGLGDKVFVNVLTELISLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLLDMSNNNLSNLQPQGFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGIMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAPAQQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTITITVKSVDTSVTIHISWKLALPMTAIRLSWLKLGHSPAEGSITET
IVTGERSEYLVTALEPDSPYKVCVPMETSNLYLFDETPVCETETAPLRMYNPTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSILFSRNCAYSKGRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCCAGAGTCCTTGCCCAGGCCACCCAGGCTTCTGGCA
GCCCTGCCGGGGCCTGTCTTCAATGTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGCTGCTGCTCTGCTTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTCAGACACTGATCCTGTCT
GCGCTGGAGAGAGGCCACCGTCTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGTGAGAAGAGCAGCTAAAAAGTGTCCGGAGAAGTGGGCCAGGAGCCCTGC
TGCAGCCGCTGAGCCTGCCGTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGTGATCCAAGTACCTAACAGAGAGTCCAGCTGACCCCTCCAGCCCG
GTTTGAAAGCTCCCACATGCCCTGGATCCACACTGATGCCCTTGGTGTACCCCACGTTGGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGTGACGTGTGCCCTGGTGCAGCTGCTGGAAACCGGGACG
GACAGCAGCGAGCCCTGCCCTCAGACCTCTGCAGGAGCCTATGACCAAGCCGGCTGCTC
AGGCTACTGCCCTGCCCACCAACTGCTCTTCTCTGGCCAGAATGAGGGATGCACACAGG
GACCACTCCAACAGAGGCCAGGACTATATCAACCTCTTCTGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATAACGCCAACCTACCCGGACATCTCATGGAAAACATCATGTTCTG
TGGAAATGGGCGGCTCTCCGACTTCTACAAGCTCCGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGAGTGAAGAGGCGAGAAAACAATTCCAGATTCTCGCTCTG
TGCTCAGGCTGGAGTACAGTGGCGCAATCTGGCTCACTGCAACCTTGCCTCTGGTTCAAGC
AATTCTTGCCTCATCCTCCGAGTAGCTGGACTACAGGAGCGTGCACCACCTGGCTAAT
TTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTATGCTGGTCTCGAACCTCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCCAAAGTGTGGATTATGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCTCCCTAGATGGCTGCTCTCCACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCCATACATCCTGGCAGAACACCCCCCAGCAAACAGAGGCCACACCCATCCACACCG
CCACCAAGCAGCCGCTGAGACGGACGGTCCATGCCAGCTGCCGGAGGAGGAACAGACCC
TTAGTCTCATCCCTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAACTTGGAGGCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTCAACTGCAACTGAAAAA
AAAAA

FIGURE 134

MSARGRWEGGGRRACRGSLGLARAQGAERVTSSEQRPMASLGLLLLLTLAPPLWSSSLPGLD
TAESKATIADLILSALERATVFLERQLPENLDGMGVVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMIDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAIIWSQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPYPLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTCTTCCCGCCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTGGAGGTGCCCGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTCAAGATTGAGGGCGTAGCTGAGTTCCAGGGTGAAGC
CTCAGGACTGGATCTGGCGGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTGGTTCTTAAG
ACAGATGGGAGTTTGTGGTCATGATATACTTCTGGATCTTATGTAGTGGAAAGTTGTATCTCC
AGCTTACAGATTTGATCCGTTGAGTGGATATCACTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAACTGCCCTATCCTCTCAAATGAAATCTCAGGT
CCACCTTCTTACTTTATTAAAAGGAATCGTGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTTCTCCTTATTGATATTGTGCTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
AAAAACAGGCAAAGTGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGATTGCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCGACGTTGATCTTACAACGTGTATGTT
AACTTTTAGCACATGTTTGACTTGGTACAGAGAAAACCCAGCTTCATCTTGCTGTAT
GAGGTCAATATTGATGTCACTGAATTAAATTACAGTGTCTATAGAAAATGCCATTAATAATTAT
ATGAACTACTATACTTATGTATTTAAACATCTTAATCCAGAAATCAAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSEVPGAAAEGSGGSVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGFLKTDGSFVVHDIPSGSYVVEVVSAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGTDFLMNPVMMMVLPLLIFVLLPKVVNTSDPDMREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGKGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCAGTTCTCCCCAGTTCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCTGGAAGGGAGGCTATATGCGTCAATTCCCCAAACAA
GTTTGACATTCCTCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGACTAACGGCGAGCCAGGATGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTCTTCACGGGAGGCTGGCAGT
TTTTCTTACTCCTGTGGTCTCAGATTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
CTTCTCTCTGCTGCGTTTATCTCTATGGACTCCTTCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTTCAGGAAATACGAAATGGATTCTGAGATACGGGCAGTG
TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCCTGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAAATGGAGGAAAGTGTGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAAGAGGAGGCATGCCCAA
ACCACCATCTCTTACTGTACTAGTCTTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
CTTCCTTGCATGATTGTCTTATGCATCCCCATCTTAATTGAGACCATACTGTATAAGATTT
TGTAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTATTTGCTATTAA
ATGTATTTATTTTACTGGACATGAAACTTAAAAAAATTCACAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTTATTTATACAGTAAAAAAAAACCTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCAATTGTATTCAACTAAGGACATATTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCACCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTTCTTGCATACCAAAAAAAA
AAAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGAAGCGCGCTGCAGCAGGGCAGGCTCCAGGTGGGGTCCGGTCCGCATCCAGCC
TAGCGTGTCCACGATGCGGCTGGCTCCGGACTTTCGCTACCTGTTGGTAGCGATCGAGGTGC
TAGGGATCGCGGCTTCCTCCTCGGGATTCTTCCCCTGGCTCCCGTTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCCAGGCCAACCCTCGCTGGAGCCAGTCTAACTGGACCACGCC
ACCACCTCTTCAGTAAAGTTATTGTTCTGATAGATGCCCTGAGAGATGATTGTGTTG
GGTCAAAGGGTGTAAATTATGCCCTACACAACCTACCTTGAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTGAATCAAGGCATTGATGACGGGAG
CCTTCCTGGTTGACGTCACTGAGAACCTCAATTCTCCTGCACTGCTGGAAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAAGCATTGTGGAATATGATGAAACAACCTCATTTTCTGTCAGATTACACAGAGT
GGATAATAATGTCACGAGGATTGGATAAAAGTATTAAAAAGAGGGAGATTGGGACATATTATCC
TCCACTACCTGGGCTGGACCACTGGCCACATTTCAGGGCCCAACAGCCCCCTGATTGGCAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCACTGCAAGGAGAGA
GACGCCCTTACCCAATTGCTGGTCTTGTTGACCATGGCATGTCAGAACAGGAAGTCACG
GGGCCTCCACCAGGAGGTGAATACACCTCTGATTTAATCAGTTCTGCGTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCATGCGGATGTGGCTGCGACACTGGCAGTAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCTATTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCACATTGAATACAGTCAGCTAGTAAACTGTTGCAAGAG
AATGTGCCGTATGAAAAAGATCCTGGTTGAGCAGTTAAATGTCAGAAAGATTGCATGG
GAACTGGATCAGACTGTAATTGGAGGAAAAGCATTCAAGCTCTATTCAACCTGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTGAAGACGCTGAGCTGTCAGCTGGCTGAGTCACAAGTGGCCAG
TTCTCACCCCTGCTCCTGTCAGCGTCCACAGGCACTGCAACAGAAAGGCTGAGCTGGAAAGTCCA
CTGTCATCTCTGGTTTCTGCTTTTATTGGTGAATCTGTTCTTCGGCGTTACGT
CATTGTGTGCACCTCAGCTGAAAGTTCTGCTACTCTGTGGCTCTCGTGGCTGGCGCAGGCT
GCCTTCTGTTACAGACTCTGGTGAACACCTGGTGTGCTGCCAAGTGTGCGACTGGCCCTGGAC
AGGGGCGTCAGGGAGGAGCTGGAGCAGCTTATCCAGGGCTCTGGGTGCCCCACACAGGTG
TTCACATCTGTGCTGTCAGATGCCCTAGTTGAGCTGAAAGCTAGGTTCTCGCACTGTTAC
CAAGGTATTGTAAGAGCTGGCGTCACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAA
TCGGACAGCCTCCAGCAGGGTGTGGAGCTGCACTGAGCTGAGGGAAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCCAGAACATGCATCCT
GCCTCATCAGTCAGATTCTCAAGCGGACGTTCTGTTGAATTCTAGTCCTGGCC
TCGGACACCTTCATTGTTAGCTGGGAGTGGTGGTAGGCAGTGAGAAGAGCGGATGGTCAC
ACTCAGATCCACAGAGCCAGGATCAAGGGACCCACTGCACTGGCAGCAGGACTGTTGGCCCC
ACCCCAACCCCTGCACAGCCCTCATCCCTCTGGCTTGAGCCGTAGAGGCCCTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTCACTCAGGGCACAGGCTCTCGGAGGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCATCTGGGCTCATGCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTTACAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTLLPPLF
SKVVIVLIDALRDDVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDHSVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFVSDYTEVDNNV
TRHLDKVLKRGDWDLILHYLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGTTATCGTGACGCACCTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGGAGTACAACCT
TCCCTATAGAAAACAACGCCAGCACCTTAAGACCACTCACACCTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTATGACCAGGATCACAAAGTACTGGTCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAAACATACGCCAGAGATCTTCTTGCAATTGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGGAGTTTGTCCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTCAGCTGAAGAAGGAGAAACTGATGAAGCT
GGCTGCCAAAAGGAATCAGCACGCCGCCCTCATCTTTATAGGGCTCAGGTGGCTCCTGGA
ACATGCTGGAGTCGGCGCTCACCCGGATGGTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGTGACAGATAAATTGAGAACAGGAAACACATTGAATTTCATTCACCAGTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCCATGAAACGCCCTCGCTA
ATTGAACTAATTGTATAAAAACACCAAAACCTGCTCACT

FIGURE 142

MLLLLEYNFPIENNQHLKTTHTFRVKNLNPKKFSIHDDQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSNMLESAAHPGWFIGTSCNCNEPVGVTDFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCCTTAA
TCCAGGATCCTGCTCCTCTGTAGGAGTGCCTGTTGCCAGTGTGGGTGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCAATGCTGGCCTGCCGTGG
AAGGGAGGTCTGCTCTGGCGCTGCTGCTGCTTCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTCCACCGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCAACAGACTGGGG
CACATCTTGAATTCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTGAAGACGACATTGACAACGCCATTCCAAGAAAGCACAG
AGCTGAACAAATCTTCACCTGCTTCTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAAG
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTTGAGTGAAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTGTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNNTFCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGGCCAAGGTTCTGACCGGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGCCACTGTCTGCATGCTCTTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAACATCAAGTGGAACCGGAAGGCCCTGCCAGCAGCTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTTCATCAAGCAAGGCCAGCAGCTGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGCATTGTACCGGCTGCATCAATGCCA
CCCAGCGGCGAACCAAGGGGAGTTCCAGAACAGCAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTGCTCCCTCAAGCATTGCGAGTTTGGAGAGGGCGCAGGACT
TCGGGTCAACCATGCACCGAGCCAGTGCTCCTCTGCCCTGGCTTGATCTGGCTCATGGTGAAAT
AAGCTTGCAGGGCTGGCACTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGGGGATGCACTCGCACTGCAA
TGCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCCTCTGATAAGATGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTCACAGTACTTCCAACAACCTTAGAGGTAG
GTGTATTCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATE
ACCAGCTAGAAAGGGCAGAGCCAGGATTCAACCCCTGGCTTGCTAACCCAGGTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTATCACTTATGCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSWWWLATVCMLLFSHLSAVQTRGIKHKRINKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNLHQQ
VLWRLVQELCSLKHCFWLERGAGLRVTMHQPVLLCLLALIWLVMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCACCATGTCAGGTCCAGAGTCATTC
CCTGATGATTTAGACTCAAAGAAAACTCATGTTCAGAAGCTCTTCTCTGGCCTCCTCT
CTGTCTTCTTCCTCTTCTTATTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCATTGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACTT CAGCACCCACAGGGCGGACAGCGCTCCCCTTACCTGGAGACTTGAC
TCCC CGCGCCCCAACCCCTGTTATCCCTGACCGTCAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCCGCCCCACACCCACCCCTCTGGCTCTTCTGTTTACTCCTCTTTCTTCATA
ACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGCTGTGACCCAAGCCGAGCGTGGAAAGAATGGGTT
CCTCGGGACCGGCACTTGGATTCTGGTGTAGTGTCCCATTCAAGCTTCCCAAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAGTGCAGAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAACATATCCTCAGAAAACAAGCCAGGTAGAGCAACTATTCTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAATTGAGAAAGAAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAAGATGTTGATTCAACCAAGAACGAAACTG
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAAATTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTAACCGCTGAAGACATTGTCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAATTAACTCAAAAGGAAGCCAAACAA
TTATGAGGAGGATCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTGCTAAGGGAGAAAAGCATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACCTCA
ATATTCCCAAATTCTATGCGCTACTGAAAAGTATTGATTCAAGAAAAGAAGCAAAGAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTGTGAAGATGATGGTGAATATGGAACAATATCT
CCAGAAGAAGGTGTTCTACCTTGAAACTTGGATGAAATGATTGCTCTTCAGACCAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCCAGCACCCTCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGGAGGAAAGACAGATGAACCCAAAGGAAAAGAGACGCTATTGGAAGC
CATCAGAAAAAAATATTGAATGGTGAAGAAACATGACAAAAGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAGGAA
GAAGCCGAGGCCATCAAGCGATTATAGCAGCCTGTAAAATGGCAAAAGATCCAGGAGTCTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAACACTTCAATTCTGTGATTAAAATTTTGACCCAAGG
GTTATTAGAAAGTGTGAATTACAGTAGTTAACCTTACAAGTGGTTAAAACATAGCTTCTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTPMAAIQDGLAKGENDETVSNTLTLNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENILDEMIALQTKNKLEKNATDNISKLFPAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCCTGTGGAGCTCAA
GATGGTCTCTGAGTGGGGCGCTGTGCTTCGAATGAAGGACTCGGCATTGAAGGGTCTTTATCTGC
ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAAGGTCAATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGTGGCTGGATGCCAGCCTGTCCCCCGTCATCTGGGTGTCAGGGTGGAAAG
CCAGTGCCTGTCATGGGGTGGGAGCCAGGAGCCACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCACAGGAATCCAAGAGCTTCACCTCTACCGCGGGACATGGGGCTCAC
TCCAGCTCGAGTCGGCTGCCTACCCGGCTGGTCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAAGACTCACCACTCCCCAGAAATGGTGGCTGGATGCCCATCACAGACTCTACTTCC
AGCAGTGTGACTTAGGGCACCGTCCCCCAGAACCTCCCTGGGCAGAGCCAGCTCGGGTGGGGGG
GAGTGGAGGGAGACCCATGGCGGACAATCACTCTCTGTCTCAGGACCCCCACGCTGACTTAG
TGGGCACCTGACCACTTGTCTGGTCTCCAGTTGGATAAAATTCTGAGATTGGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGCCTACTGCTGTGGAACCTTGTAAAAAACCATGTGGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGGAGTGGTGGGAATCATTCTGCT
TAATGGTAACGTACAAGTGTACCTGAGCCCCGCAAGGCCAACCCATCCCCAGTGAGCCTTATA
GGGTCACTAGCTCTCCACATGAAGTCTGTCACTCACCAGTGTGCAGGAGAGGGAGGTGGTATA
GAGTCAGGGATCTATGGCCCTTGGCCCAGCCCCACCCCTTCAATTCTGCCACTGTCTATA
TGCTACCTTCCTATCTTCCCTCATCATTTGTGTCGGCATGAGGAGGTGGTGTAGTCAGAA
GAAATGGCTCGAGCTAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTTTAAAAACCCAA
GATACAATCAAATCCCAGATGCTGGTCTCTATTCCATGAAAAAGTGCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATAATTGCAATTATTAAATTAAAGATAACCTATTATATT
TCTTTATAGAAAAAGTCTGAAAGAGTTACTCAATTGTAGCAATGTCAGGGTGGCAGTAT
AGGTGATTTCTTTAATTCTGTTAATTATCTGTATTCCTAATTTCTACAATGAAGATGAA
ATTCTCTGTATAAAAATAAGAAAAGAAATTATCTTGAGGTAAAGCAGAGCAGACATCATCTGAA
TTGTCCTCAGGCCCTCACTTCCCCAGAGTAAATTCAATTGAAATCGAGCTCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGGCCATGAGGAGGGCTGTGCTGAGTTGT
GTGGCTGGAATCTGGGTAAGGAACCTTAAAGAACAAAATCATCTGTAATTCTTCTAGAAG
GATCACAGCCCCCTGGGATTCAGGCAAGGCATTGGATCCTCAAGAAGGCTGCTGACTGGTTGA
ATTGTGTCCTCCCTCAAAATTCAACATCCTCTTGGAAATCTCAGTGTGAGTTATTGGAGATAAG
GTCTCTGAGATGTAGTTAGTTAAAGACAAGGTATGCTGGATGAAGGTAGACCTAAATTCAATT
GACTGGTTCTCTGTATGAAAGGAGAGGACACAGAGACAGAGGAGACGGGGAAAGACTATGTA
AAGATGAAGGCAAGAGATGGAGTTTGCAAGCCACAAGCTAAGAAACCCAAGGATTGGCAACC
ATCAGAAAGCTTGGAAAGGCAAAGAAGAATTCTCCCTAGAGGCTTAGAGGGATAACAGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAAGAAAGAATAATTCTGGCTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTTAGGAAACTAATACAGTGTCTAAATGATCCCTGT
CTCCTCTGTTTACATTCTGTGTGTCCTCCACAAATGTACCAAGATTAGTTATAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTCTGCCCCACCCACGGCCCCAATCTATCTGGCTACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGGCTATAAAGAGACTTACGTGTAAGGGGGGGGGGGGG
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTGGTTTT
AAGTTGCTCAGTTGGTCTAATTGTTATGCAGCAATAGATAAAATAATGCAAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEIISVVPNRWLDASLSPVILGVQGGS
QCLSCGVGQEP TLLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTCCTTATGGGGACCCCTGCCACCAGCTGCCTCCTCTCTGG
CCCTCTGGTACAGGGAGGAGCAGCTGCCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTGCTCATTGGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCTTGAAGAAGTGTGTTCCCTCAATCTGATAGGTTC
CAGCCTTATATGCAAGGAGGTGGTGCCTGCCCTCCTGCCAGGCTCAGCAACAGGCTAACGACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCAAAAGCGATTTTTAACCAAAAGGAAGATGGAAGCCAAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTAGTTACAAAGGAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTGATAACATTGTAACTGGTGTTC
TATACACAGAAAACAATTATTTAAATAATTGTCTTTCCATAAAAAAGATTACTTCCAT
TCCTTAGGGGAAAAACCCCTAAAGCTCATGTTCCATAATCAGTACTTATATTATAAA
TGTATTATTATTATTATAAGACTGCATTTATTTATATCATTATTAAATATGGATTATTAT
AGAAACATCATTGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKSVSSFIMGTLATSCLLLALLVQGGAAAPISHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVRLIGEKLFHGVSMERCYLMKQVLNFTLEEVLFQSDRFQPYMQEVVPMFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELLLLMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTCCCCGACTGTGACTGAGTGCGAGTGCAGTCCCAGCATGTACAGGTCAGTCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCAAGC
TGCCAGGTTGGGCTGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTCCTACAGGTGGTTGCAT
TCTTGCAATGGTCATGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCCTCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTAACAGCAGGCCATCT
CCCCCTGGAGATATGAGTTGACAGAGACTGAAACCGCTCCACATGGACCCCCGGGCAACTCGGA
TGCCCTGTGCCCGACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTACAGGCGCCATGCCATGGCGAGAAGGGCACCA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTAGCTTGTGTGTGCGGGCCCGT
GTGATGGGCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCATGAAGGGCCAGGATGCCAGATGCTTGGCCCTGTGAAGTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGCTTGGGAAAACCTGCACTCTGCACATTTGAAAAGAG
CAGCTGCTGCTAGGGCCGCCGAAGCTGGTGTCCCTGTCACTTCTCTCAGGAAAGGTTTCAA
GTTCTGCCATTTCTGGAGGCCACCACTCTGTCTTCCCTTTCCATCCCTGCTACCTG
GCCAGCACAGGCACTTCTAGATACTTCCCTGCTGGAGAAGAAAGGCCCTGGTTTATT
TGTTTGTAACTCATCACTCAGTGAGCATCTACTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTGTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTATCAAATAAATAT
CTTTATTTAAAAATGAAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPKGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**A**TGTCGCTCGTGCTGCTAAGCCTGGCCGCCTGTGCAGGAGCGCCGTACCCGAGAGCC
GACC GTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGACTATTCA
ATTTGATGAATGTAAGCTGGTACTCCGGCAGATGCCAGCATCCGTTGTTGAAGGCCACCAA
GATTTGTTGACGGGAAAAGCAACTTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTGGTGGTAAATGGACATTTCCTACATCGGCTCCCTGTA
GAGCTGAACACAGTCTATTCATTGGGCCATAATATTCCATAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTCACCTCACCAAGGCTGCCAGACACATAATGAAATATAAAAAAA
AGTGTGCAAGGCCGAAGCCTGTGGGATCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
GTAGAAAGTGAACCTCACAAACCACTCCCCTGGAAACAGATACTGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTCTCAGGTGTTGAGCCACACCAGAAAGAAACAAACGCGAGCTCAGTGGTGA
TTCCAGTGACTIONGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTCCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTGCCACAAACAGGCGTCCCTTCCCT
GGATAACAAACAAAAGCAAGCCGGAGGCTGGCTGCCTCTCCTGCTGTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAAATGTGGAGGCACGAAAGGATCAAGAAGACTCCCTT
TCTACCACCAACTACTGCCCCCCATTAAGGTTCTGTGGTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTTCACTGAATTCTCAAAACATTGCAAGTGAAGTGGCTCATCCTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGTCCAGTGCAGTGGCTGCCACTCAAAGAAGGCA
GCAGACAAAGTCGCTTCCTCTTCCAATGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTCCCCCTGCCCTTAACCTTCTGCA
GTGATCTAAGAAGCCAGATTCATCTGCACAAATACGTGGTGGTCACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCACTCATGAAGGATGCCACTGCTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCCCTGT**AG**

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYSILMNWSW
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEGDPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVFPFLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVYPSEICFHHTICYFTEFL
QNHCRCSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-492

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCC~~T~~GATGCCCATGGTCAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTCTGAGTGAGGCGCAGCTCGGAAATCCCCAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCTTCCATGTCACTGTAACATCGAGAGCCGCTCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCACCGGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGA~~CT~~GTTGGCTGCACCTGCGTCACCCCTGT~~CAT~~CCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 16o

MTVKTLHGPAMVKYLLSILGLAFLSEAARKIPKGHTFFQKPESCPPPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
ACAGAGAGTGACAAAATACCCAGCACAGCCCCCTCGCCCCCTCTGGAGGCTGAAGAGGGATT
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTGGGGGGGGCAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACTAGAAGATGCCTGTGCCCTGGTCTTGCTGCCT
TGGCACTGGGCCGAAGCCCAGTGGTCTTCTCTGGAGGAGCTGTGGGGCCTCAGGACGCTACC
CACTGCTCTCGGGCCTCTCTGCCGCCTGGGACAGTGACATACTCTGCCTGCCTGGGACAT
CGTGCTCTCGGGCCCTGCTGGCCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTGGCGTGCATGGGACTGG
GAAGAGCTGAAGATGAGGAAAGTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
TGCCCTCTCCAGGCCAAGTCGTGCTCTCCAGGCCTACCCACTGCCCCGTGCGTCTGC
TGGAGGTGCAAGTGCCTGCTGCCCTTGTGCAGTTGGTCAGTCTGTGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGCTCTACTCAGCCAGGTACGAGAA
GGAACTCAACCACACACAGCAGCTGCCCTGCCCTGGCTCAACGTGTCAGCAGATGGTACA
ACGTGCACTGGTCTGAATGTCCTGAGGAGCAGCACTCGGCCCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCAAAACCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTGAA
CCACACAGACCTGGTCCCTGCCCTGTATTAGGTGTGGCCTCTGAAACCTGACTCCGTTAGGA
CGAACATCTGCCCTTCAGGGAGGACCCCCCGCAGACACCAGAACCTCTGGCAAGCCGCCGACTG
CGACTGCTGACCTGCAGAGCTGGCTGCTGGACGCACCGTGCCTGCCGCAGAACGGCACT
GTGCTGGGGCTCCGGTGGGACCCCTGCCAGCCACTGGTCCCACCGCTTCCCTGGGAGAACG
TCACTGTGACAAGGTTCTCGAGTTCCCATTTGCTGAAAGGCCACCTAACCTCTGTGTTAGGTG
AACAGCTGGAGAAGCTGCAGCTGCAGGAGTGCTTGTTGGCTGACTCCCTGGGCTCTCAAAGA
CGATGTGACTGTGGAGACAGCAGGGCCCCCAGGACAACAGATCCCTCTGTGCCCTGGAAACCA
GTGGCTGACTTCACTACCCAGCAAAAGCTTCAAGGGCAGCTGCCCTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTGCACTATGGACGTGACTTGGGAGCGCTATGGGCTG
CCCCATGGACAAATACATCCACAAGCGCTGGGCTCTGTGTGGCTGGCCTGCCTACTTTGCCG
CTGCGCTTCCCTCATCCTCTCAAAAGGATACCGCGAAAGGGTGGCTGAGGCTCTTGAAA
CAGGACGTCCGCTGGGGCGGCCAGGGGCCCGGGCTGTGCCCTACTCAGCCGATGA
CTCGGGTTTCGAGCGCTGGTGGCGCCCTGGCGTCGGCCCTGTGCCAGCTGCCCTGCCGTGG
CCGTAGACCTGTGGAGCCGTGTAACGTAGCGCGCAGGGGCCGTGGCTTACGCCAG
CGGCCAGACCCCTGCAGGAGGGCGCGTGGTGGCTTGCTCTCTCCCGGTGCCGTGGCCT
GTGAGCGAGTGGTACAGGATGGGTGTCGGGCCGGGGCGCACGGCCCGCACGACGCCCTCC
GCGCCTCGCTCAGCTGCGTGTGCCGACTTCTTGCAAGGGCCGGCGCCCGCAGCTACGTGGGG
GCCTGCTTCGACAGGCTGCTCACCCGGACGCCGTACCGCCCTTTCGCACCGTGCCGTCTT
CACACTGCCCTCCAACGTGCCAGACTTCTGGGGCCCTGCAAGCAGCCTGCCCGCTCCG
GGCGGCCAGAGAGAGCGGGAGCAAGTGTCCCAGGGCCCTCAGCAGCCCTGGATAGCTACTTC
CATCCCCCGGGACTCCCGCCGGGACGCCGGGGTGGGACCGGGGACCTGGGGCGGGGA
CGGGACTTAAATAAAGGCAGACGCTGTTTCTAAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVGPQDATHCSPGLSCLRWDSDLCLPGDIVPAPGPVLAPTHLQTELVLRCQKETDCDLCLRVAVHLAVGHWEEPEEKFQGGAADSGVEEPRNASLQAQVVLFSFQAYPTARCVLLEVQVPAALVQFGQSVGSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSADGDNVHLVLNVSEEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDHSVRTNICPFREDPRAHQNLWQAARLRLLLTQSWLADAPCSLPAEAALCWRAPIGGDPCQPLVPPLSWENVTVDKLEFPPLLKGHPNLCVQVNSSEKLQLQECLWADSLGPLKDVLLETRGPQDNRSILCALEPSGCTSLSKASTRAARLGEYILLQDLQSQCQLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLLLLKKDHAKGWLRLLKQDVRSGAAARGRAALLLYSADDSGFERLVLGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRRQTLQEGGVVLLFSPGAVALCSEWLQDGVSAGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTPVFTLPSQLPDFLGALQQPRAPRSGRLQERAEVQSRALQPALDSYFHPPGTPAPGRGVPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251, 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617, 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTGACTGTGGGATCCCTGGCT
GCTCACGCCCCGAGGACCCCTCGGATCTGCTCCAGCAGTGAATTCCAGTCAGCAACTTTGA
AAACATCCTGACGTGGGACAGCAGGGCAGAGGGCACCCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGTGTGCAGGGATCACCCGAAGTCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGGTACCGC
GTCAGTGCGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTCACTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGATCTCAAAGTGAGATGATTCACTGAGATGATTGTTCATCCTA
CCCCCACGCCATCCGTGCAGGCATGGCACCCTGAACCTGGAAAGACATCTTCCATGACCTG
TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGACCTTGGAGGGAAAGCAGAGAGA
ATATGAGTTCTCGGCCTGACCCCTGACACAGAGTCTGGCACCATCATGATTGCGTCCCCA
CCTGGGCCAAGGAGGTGCCCCCTACATGTGCGAGTGAAGACACTGCCAGACGGACATGGACC
TAATCTCTCGGAGGCCTCTGTTCTCCATGGGCTCTCGTCAGTACTCTGCTACCTGAG
CTACAGATACTGACCAAGGCCCTGCACCTCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
TCCAGCCGTGCGCTTCACTCAGGAGCACGTCTGATCCCTGTCTTGACCTCAGCGGCCCGC
AGTCTGGCCAGCCTGTCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGGCCGCAGGAGC
TCCACAGCGGCTAGCCTGTCAGTACCTACTTAGGGCAGCAGACATCTCATCCTCCAGC
CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCCTATGCCCAAACGCTGCCCTGAG
GTCGGGCCCCATCCTATGACCTCAGGTGACCCCCGAAGCTCAATTCCATTCTACGCCCA
GGCCATCTCTAAGGTCAGCCTCTCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
CCTATGGGGTATGCACTGGAAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCTAA
CACCTTAGGCTAAAGGTCAAGCTTCAGAAAAGAGCCACAGCTGGAGCTGCATTTAGGTGCC
TTCTCTGCAGGAGGTGACCTCTGGCTATGGAGGAATCCAAGAAGCAAATATTGCA
CCCTGGGATTGACAGACAGAACATGACCCAAATGTGCTACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCCAGCTCCCCCTCTCTCAGTCCAGATGAGGGCCACCCATGTC
CCTCCCCCTTGCAACCTCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCAAGTCCCTGGGCC
TGCTGGAGTCCCTGTGTCTCCAAAGGATGAAGCCAAGGCCAGCCCCCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACACTGGATTCTCTTTTCAAGGGCTGGCCCTGACTGTGCA
AGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGGGGGAGCTCTGGGAGCAGCTGTGTAGACAAGCGGTGCTCGTGAGCCCTG
CAAGGAGAAATGACAGTGCAGGAGGAATGCAGGGAAACTCCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGTCAAGGAAATTGCCCTCTCTGTCAATTGTTCAAAGGTGG
ACAATCTAGCTGACAGAGCATGAGGCCCTGCCTCTCTGTCAATTGTTCAAAGGTGG
GCCCTGGAAAAAGAACCAAGCCTGGAAAAGAACCAAGAGGAGGTGGCAGAACAGAACACTGC
ACTTCTGCCAAGGGCAGGGCAGCAGGACGGACTCTAGGGAGGGTGTGGCCTGCAGCTCA
TTCCCAGGCCAGGGCAACTGCCCTGACGTTGACGATTTCAGCTTCAATTCTCTGATAGAACAAAGC
GAAATGCGAGTCCACCAGGGAGGGAGACACACAAGCTTTCTGCAAGGAGGTTCAGACCC
ATCCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
TGTACTGATGTCACAACCTTGCAAGCTGCTGCCCTGGGTCAGGCCATCTGGCTCAAATTCCAGC
CTCACCACTCAAGCTGTGTACTTCAAACACCTACCTCATGGAGTTGTGGTAAGATGAAAGT
ATCTGTAATGTTGGGATCATAACACCTACCTCATGGAGTTGTGGTAAGATGAAAGT
TCTTTAAAGTGTCTTAATGTCCTGGTACATGGCAGTGCCTAAACGGTAGCTATTAAAAAA
AAAAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWSGPETPDTVYSIEYKTYGERDW
VAKKGCCRIRKSCNLTVETGNLTELYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVT CIS
KVRSIQMIVHPTPTPIRAGDGHLTLLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNQQRVLTFQPLRFIQEHVLIPVFDLSPGSSLAQPVQYSQIRVSGPREPAGAPQRHSLEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLLSPKHLRPKGQLQKEPPAGSCMLGGSLQEVTS LAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESLVCPKDEAKSPAPETS DLEQPTELDSLFRGLALT VQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCGCGGTGCCACAACATGG
CTGGCGGCCGGGCTGCTCTGGCTGTTCTGGACCTCAAAGTGTGGGGACGAAGAGTCAGCAT
GTTAATGTACCGTGGGAAAGCTCTGAAGACTTCACGGCCCTGATTGTCGTTGTGAATTAA
AAAAAGGTGACGATGTATATGCTACTACAAACTGGCAGGGGATCCCTGAACCTTGGCTGGA
AGTGTGAACACAGTTGGATTTCAAAAAGATTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTGTCTGCTTGAAAGGAGGAAGAGATGATT
TTAATAGTTATAATGAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCTGTGCCTGAGCCGAGGCATTAGAGCTGATTAGAGGATGGAGAAGGTGCTTCT
CAGAGAGCACCGAGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTCGTTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGAAGCGAAAGCAGAAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTCGTTGGCATAAAATCTAAGTTGTTACAAAGATTGTTAGTA
CTAAGCTGCCCTGGCAGTTGCATTTGAGCAAACAAAATATTATTTCCCTCTAAGTA
AAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFVLGALWWVPGQSDLSHGRRFSDLKVCDEECMLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPADETDFVCFEGGRD
DFNSYNVEELLGSLELEDSPVEESKKAAEVSQHREKSPPEESRGRELDPVPEPEAFRADSEDEGA
FSESTEGLQGQPSAQESHPTHSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCIVIHYSKGFRWHQNLSFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAAGAGGCCGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCAAAGCCGGGCTTCGCTTCGCCTCTAGGACATACACGGGACCCCCCTAACCTTC
AGTCCCCCAAACGCGCACCTCGAAGTCTTGAACTCCAGCCCGCACATCCACGCCGGCACAGG
CGCGGCAGGCGCAGGTCCCAGCGAAGGCAGTCGCGCAGGGGTGCGGCAGCTGGCTCGGC
GGCGGGAGTAGGGCCCGCAGGGAGGCAGGGAGGCTGCATATTCAAGAGTCGCGGGCTGCCTG
GGCAGAGGCCGCCCTCGCTCACGCAACACCTGCTGCTGCCACCGCGCCGCATGAGCCCGTGG
TCTCGCTGCTGCTGGCGCCCGCTGCTCGGCCACGGAGCCTCTGCCGCCGCGTGGTCAGC
GCCAAAAAGGTGTGTTTGCTGACTTCAGCAGTCAAGCATCCCTGCTACAAAATGGCTACTTCCATGAAC
GTCCAGCCGAGTGAGCTTCAGGAGGCACGCCCTGGCTGTGAGAGTGAGGGAGGAGTCCTCTCA
GCCTTGAGAATGAAGCAGAACAGAACAGTTAATAGAGACATGTTGCAAAACCTGACAAAACCGGG
ACAGGGATTCGATGGTGATTCTGGATAGGGCTTGGAGGAATGGAGATGGCAAACATCTGG
TGCCTGCCAGATCTTACCACTGGTCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTCCCTGCGGAAGTGAAAAGTGTGTTGATGTATCACCAACCAACTGCCAATCTGGC
CTTGGGGTCCCTACCTTACCACTGGAAATGATGACAGGTGTAACATGAAGCACAATTATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTTACTGAAGCAGGTATAATTCCAATCTAATTATGTGTT
ATACCAACAATACCCCTGCTCTACTGATACTGGTTGTTGGAACCTGTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAAACTAGTCAAACAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAAGTGGCATGAATAAACTCATTGACTTGGTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTTGAATGGCTGAAATCAAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTGAGGCAAATATTAAAGTAATTTTATATGTCTATTATTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTGCTAAAGGATGCACCAA
ACTTCAAACCTCAAGCAAATGGAATGCAAGATAAAGTGTATCAACACGTGGAGTA
TGTGTGTTAGAAGCAATTCTTATTCTTCACCTTCTGTTGTTGTTGTTGTTGTTGTTGTT
TGTATATTGTTGAAATTACAGTGTGCAAAGTATTTCACCTTGTGTTGTTGTTGATAAAA
ATGAACTGTTCTAATATTATTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT
AAACTTATTACTGTGTCACGAAATTCAACACACACAAATAGTACCATAGAAAAAGTTGT
TTTCTCGAAATAATTCTATCTTCAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT
AATAAGAAGCTATTCTATTAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT
TGCTAATTCAATTGTCAGACATGTGCTTATAATTATTATTGCTTAAATTAAACAGATT
TTGTAATAATGTAATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT
TGACATACACAATATAATCATATGTCTTCACACGTTGCTTATATAATGAGAACAGCTCTGA
GGGTTCTGAAATCAATGTGGCCCTCTTGCCCACTAAACAAAGATGGTTGTTGTTGTTGTTGTT
ATTGACACTGGAGGGCAGATAGTTGCAAAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT
ACTATATTAGTATACAAAGAGGTCTATGTGGTTGAGACCAAGGTGAATAGTCACTATCAGTGGAG
ACAAGCACAGCACACAGACATTAGGAAGGAAGGAACATACGAAATCGTGTGAAATGGTTGG
AACCCATCAGTGTGATCGCATATTCAATTGATGAGGGTTGCTTGTGAGATAGAAAATGGTGGCTCTT
CTGTCTTATCTCTAGTTCTCAATGCTTACGCTTGTGTTCTCTCAAGAGAAAGTGTAACTCT
CTGGTCTTCAATATGCTTGTGCTTCAACAAATAAAAGAGTCTGTGTTCTGGGGAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVSLLGAALLCGHAFCRVVSGQKVCADFHKPCYKMAFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGTSGACPDLYQWSDGSNSQ
YRNWTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217